Alma Mater Studiorum - Università di Bologna

DOTTORATO DI RICERCA IN

MECCANICA E SCIENZE AVANZATE DELL'INGEGNERIA

Ciclo 33

Settore Concorsuale: 09/C2 - FISICA TECNICA E INGEGNERIA NUCLEARE

Settore Scientifico Disciplinare: ING-IND/19 - IMPIANTI NUCLEARI

MATHEMATICAL AND NUMERICAL MODELS FOR BOUNDARY OPTIMAL
CONTROL PROBLEMS APPLIED TO FLUID-STRUCTURE INTERACTION

Presentata da: Andrea Chierici

Coordinatore Dottorato

Marco Carricato

Supervisore

Sandro Manservisi

Co-supervisore

Ruben Scardovelli

Esame finale anno 2021

Contents

List of Figures

Bibliography

Introduction

The main purpose of this work is to develop numerical methods for the optimal control of fluid-structure interaction simulations. In particular, we focus on the Koiter shell model and on the adjoint formalism for control problems. Moreover, fractional operators are introduced to be applied to the framework described above.

In recent years numerical simulations of fluid-structure interaction (FSI) problems have gained popularity and interest in the research community thanks to the great variety of possible applications, ranging from wind turbines and aircraft to hemodynamics. In FSI problems the fluid flow changes the stress state of the solid structure that is left free to move while the solid deformation affect the fluid flow motion. A large variety of books and articles about different fluid-structure models have been published and the interested reader can consult [1, 2, 3].

Whenever many solids have small dimension in comparison with the fluid domain, various techniques have been developed to reduce the computational cost of FSI problems. Very interesting models are based on the reduction of dimensionality of the solid through the Koiter shell equations [4]. To couple the fluid and the structure domains, the Koiter shell equations are embedded into the fluid motion through boundary equations that may be treated implicitly, so that the stability of the numerical fluid scheme is preserved [5]. This model can be applied in many cases where a fluid interacts with a thin membrane that deforms mainly in the normal direction.

In the first Chapter the basics of the fluid-structure interaction theory

is introduced. In particular, a monolithic approach is studied with different models for hyperelastic structures. First we report numerical benchmarkings of such a model that have been carried out in a previous work [6]. Then, the mathematical and numerical theory for Koiter fluid-structure interaction is reported. We also report a mathematical introduction to absorbing boundary conditions, that are of great importance when pressure waves are involved, e.g. mainly in hemodynamic applications. Both the Koiter's shell model and the absorbing boundary conditions are tested and compared with results available in the literature.

In the second Chapter, we introduce an adjoint-based optimal control theory of the Koiter's fluid-structure model in the steady case by using a Lagrange multiplier approach. Many works on the optimal control of Navier-Stokes systems have been published (see, for example, [7, 8, 9]) but only a few studies on the optimal control of FSI systems can be found in literature, see [10, 11]. Using this multi-scale Koiter model, we study the FSI steady optimal control problem with mixed boundary conditions and investigate some theoretical aspects of the optimal solution.

The mathematical formulation of a simple control problem is first presented. Then, the boundary optimal control theory is applied to the fluid-structure Koiter's model for the study of solution existence, regularity and differentiability properties of the fluid-structure problem. Since the mathematical problem is not straightforward, some hypotheses are added to the Koiter's model to prove all cited properties. In the same chapter some numerical results are shown, obtained through an iterative numerical algorithm that minimizes the objective functional.

In the last Chapter, the fractional operators are introduced. The numerical implementation of operators in fractional Sobolev spaces is a topic of increasing interest and many works have been published recently on this subject [12, 13]. In particular, we apply the fractional calculus to the optimal control problems introduced in the second Chapter. The objective functionals of many optimal control algorithm problems depend on a regularization term on the controlled variable. The control parameter often belongs to a fractional Sobolev space (e.g. the restriction of a variable in the Sobolev space H^1 to the boundary belongs to $H^{1/2}$), so the proper operator to be used in such spaces is a fractional one.

The numerical implementation of the fractional operators (and, in particular, the fractional Laplacian) is not easy due to the nonlocal behavior of the

operator. Therefore, in this Chapter some numerical techniques are presented in order to simulate properly the fractional Laplacian. These are based on direct implementation of the nonlocal operator [14, 15] or on spectral theories applied to the fractional calculus [16, 17]. Then, some applications to simple control problems are reported, together with some preliminary results.

All the numerical simulations presented in this work, with the exception of the fractional simulations presented in the last Chapter, have been simulated with the in-house multigrid finite element based code FEMuS. The code is integrated into a numerical platform where various different codes are integrated to develop multi-physic and multi-scale simulations. The source code of FEMuS on GitHub can be found in [18], and the numeric platform can be found in [19]. Some works in literature introduce the code and the numerical platform, see, e.g. [20, 21]. The results of the last Chapter on the fractional operators have been carried out with the FEMTTU code, developed at the Department of Mathematics and Statistics of the Texas Tech University [22].

CHAPTER 1

Fluid-structure interaction models

Fluid-structure interaction (FSI) is a set of numerical problems where one or more solid structures deform due to interaction with a fluid that surrounds the solid themselves. The fluid flow depends on the shape of the structure and its motion, so a mutual dependence between the two regions is established. Fluid-structure interaction problems are common in many fields, like engineering, science, and medicine. Software projects of both open-source and commercial type have been developed in order to perform numerical simulations of FSI phenomena. Such simulations are becoming of great interest in the scientific community since they play an important role in the design of several components. For example, aircraft wings, airport windsock, wind-turbine blades, automobile airbags are all components that are usually modeled and designed by using the FSI models. Moreover, a lot of different natural phenomena can be simulated with such a technique, like the falling of a leaf, the dynamics of parachutes and the blood flow inside the blood vessels of various types. There are many works in literature on this topic, concerning all these applications. The interested reader can see [1, 3, 2, 23, 24, 25] and reference therein.

Fluid-Structure Interaction problems are characterized by an intrinsic mathematical challenge due to the inherent nonlinearity of the domain that

moves as a function of the physical variables. For this reason the analytical solution of such equations is extremely difficult to find, and only a few cases have been studied analytically, where simplifying assumptions have been invoked to arrive at closed-form solutions of the underlying partial differential equations. Therefore, the effort of many research groups is targeted to the development of numerical techniques for such problems.

The fluid and the solid equations, together with their respective boundary conditions, must be satisfied simultaneously. The two systems are coupled at the fluid-structure interface, which requires a set of interface conditions, i.e. are the kinematics compatibility laws and stress balances on the considered interface. As the structure moves through space, the shape of the fluid subdomain changes to conform to the motion of the structure. This is known as the Lagrangian description of the structural motion. One of the challenges of the FSI simulations is related to the fact that the motion of the fluid domain needs to be accounted in the differential equations and in the boundary conditions. There are two major classes of methods for tracking the structure deformation, which are known in the discrete setting as the non-moving grid and moving grid approaches. Furthermore, the motion of the fluid domain is not known a priori since it is a function of the unknown structural displacement. This makes FSI a three-field problem, where the third unknown is the motion of the fluid domain.

Many works on fluid-structure interaction simulations are based on a mesh separation between the fluid and the solid domains. This represents a flexible approach but the imposition of the conditions on the interface is not straightforward, in particular when the meshes of the two sub-problems don't match on the interface. On the other hand, many algorithms work on a monolithic system, composed by the fluid and the solid equations, and the interface conditions. Therefore we can define two major classes of FSI coupling techniques: loosely-coupled and strongly-coupled, which are also referred to as staggered and monolithic, respectively.

In loosely-coupled approaches, the equations of fluid mechanics, structural mechanics and mesh moving are solved sequentially. For a given time step, a common loosely-coupled algorithm requires the solution of the fluid mechanic equations with velocity boundary conditions, coming from the structure displacement at the interface, followed by the solution of the structural mechanic equations with the updated fluid mechanics interface traction, and followed by the solution of the mesh moving equations with the updated structural dis-

placement at the interface. This allows the use of existing fluid and structure solvers, instead of creating a new one. However, convergence difficulties are encountered sometimes, most commonly when the structure is thin and the fluid is heavy, and when an incompressible fluid is fully enclosed by the structure. In strongly-coupled approaches, as introduced above, the equations of fluid, structure and mesh moving are solved simultaneously. The main advantage is that strongly-coupled solvers show fewer convergence issues. However, strongly-coupled approaches necessitate writing a fully-integrated FSI solver, virtually precluding the use of existing fluid and structure solvers. Using this approach, iterations are performed within a time step to simultaneously converge the solutions of all the equations involved. In addition, several choices are possible for the discretization procedure (decouple then discretize or vice-versa), the order between the discretization and the linearization procedures (linearize then discretize or vice versa), the linearization scheme (fixed-point, relaxed fixed-point, quasi-Newton, Newton), and the choice of the linear solvers and pre-conditioners. Another computational challenge in some FSI applications involves very large structural displacements. In this case, a robust mesh moving technique is needed and the option to periodically regenerate the fluid mesh (i.e., re-mesh) to preserve the mesh quality and consequently the accuracy of the FSI computations. The re-meshing procedure requires the interpolation of the solution from the old mesh to the new one. However, the re-meshing technique can be very expensive from the computational point of view. A different approach to the large displacement cases is the immersed boundary (IB) approach, where no re-meshing is needed. In recent years the interest in the IB techniques applied to the FSI numerical simulations is growing and some works are available in literature.

In the first part of this chapter a monolithic approach with quasi-direct coupling will be presented and used [26, 27] for the solution of the FSI system. The fluid, structure, and mesh systems are treated as two separate blocks, and the nonlinear iterations are carried out one block at a time. In the second part of this chapter, a shell model based on the Koiter shell theory will be introduced, in order to reduce the computational cost of the numerical simulations involving thin structures. In fact, the shell model for the structure allows reducing by one the dimension of the solid structure, and to treat the solution of the FSI problem as a fluid problem with a moving mesh and a Robin boundary condition [4, 5].

As introduced before, in this work all the numerical simulations, with the exception of the fractional operators presented in the last chapter, are developed with the code FEMuS [18]. FEMuS is a finite element based in-house code built with hierarchical C++ classes. Several solvers have been developed until now, including solvers for the solution of the Navier-Stokes equations, turbulence, fluid-structure interaction, etc.

FEMuS is integrated into a computational platform [28], used and developed in this thesis, based on a C++ main program that handles several external open source libraries, such as the libmesh and PETSc libraries. Libmesh is a C++ finite element library used in our computational platform to generate and handle the numerical mesh and to refine it with multiple levels [29]. PETSc is a C++ library for linear and non-linear algebra developed using LASPack cores written in Fortran and other solvers. A key feature of this library is its ability to handle parallel solutions of systems of equations with solvers designed to scale very well with the increasing number of nodes by defining parallel vectors and matrices classes [30, 31].

1.1   Kinematics and conservation laws

In this section, an introduction to the governing equations of the FSI problems is presented. In particular, we will focus on the kinematics of the continuum, which is the branch of mechanics that studies the motion properties, such as position and velocity. Let L be a characteristic dimension (e.g. a side of a cube or the radius of a sphere) of a certain solid object $\mathcal{O}$. We can define $\mathcal{O}$ as a continuum media if $\lambda \ll L$, where λ is the mean free path of the particles of $\mathcal{O}$. Then the conservation laws for this particular physical problem are presented. In the last part of the section, basic models for both structure and fluid are introduced. In this section, we will follow the formalism of [6].

1.1.1   Kinematics of the continuum

Let $\hat{\Omega} \subset \mathbb{R}^3$ be a reference solid structure configuration and $\hat{\Gamma}$ be its boundary. We now consider the evolution over the time of $\hat{\Omega}$, and in particular let $\Omega \subset \mathbb{R}^3$ be the configuration at a certain time t, and Γ its boundary. In the following we consider $\hat{\Omega}$ as the configuration at the initial time $t = 0$. We define now a mapping between $\hat{\Omega}$ and Ω.

Definition 1.1. *Let $\hat{\boldsymbol{x}}$ be the position of an arbitrary point in the reference configuration and $\boldsymbol{\eta}$ be the displacement of the material point with respect to the initial configuration. Following the introduced notation, we can now define a mapping $\mathcal{X}$ such as*

$$\boldsymbol{x} = \mathcal{X}(\hat{\boldsymbol{x}}, t) = \hat{\boldsymbol{x}} + \boldsymbol{\eta}(\hat{\boldsymbol{x}}, t), \tag{1.1}$$

which maps the coordinates $(\hat{\boldsymbol{x}}, t)$ of material points in the reference configuration $\hat{\Omega}$ to their counterparts in the current configuration Ω. The $\mathcal{X}$ map is a morphism between algebraic varieties, also called regular map, which implies that is bijective, and then monodromic and invertible. Therefore

$$\forall \hat{\boldsymbol{x}} \in \hat{\Omega}, \; \exists! \, \boldsymbol{x} = \mathcal{X}(\hat{\boldsymbol{x}}, t), \, \boldsymbol{x} \in \Omega \tag{1.2}$$

and vice versa, so that

$$\hat{\boldsymbol{x}} = \mathcal{X}^{-1}(\boldsymbol{x}). \tag{1.3}$$

We call the set $(\hat{\boldsymbol{x}}, t)$ Lagrangian coordinates (called also material or reference coordinates). The Lagrangian description is well suited to describe the motion of a body defined as a fixed set of material particles. The body can change its shape under the action of external or internal forces but not its composition. On the other hand, the set $(\boldsymbol{x}, t)$ is called Eulerian or space coordinates. The application of the Eulerian coordinates is the basis of the Eulerian description of the space, which is well suited to describe the fluid flow through a fixed spatial region. In this case, the fluid particles enter and leave the considered domain and are taken into account only during transit inside the domain.

The transformation properties

Since the transformation $\mathcal{X}$ underlies all the fluid structure interaction mathematical introduction, we now analyze it and some of its properties. For this purpose the deformation gradient is now introduced.

Definition 1.2. *We define the deformation gradient $\boldsymbol{F} : \hat{\Omega} \to \mathbb{R}^{3 \times 3}$ as*

$$\boldsymbol{F} = \nabla \mathcal{X}(\hat{\boldsymbol{x}}, t) = \frac{\partial \hat{x}}{\partial \hat{\boldsymbol{x}}}, \tag{1.4}$$

and, considering (1.1) *we have*

$$\boldsymbol{F} = \boldsymbol{I} + \frac{\partial \boldsymbol{\eta}}{\partial \hat{\boldsymbol{x}}}. \tag{1.5}$$

Clearly, we can define an inverse deformation gradient $\boldsymbol{G}$ of the inverse mapping relating $\hat{\Omega}$ to the deformed configuration Ω. We can determine a relation between the two introduced operators, and in particular

$$\boldsymbol{G}(\boldsymbol{x},t) = \boldsymbol{F}^{-1}(\hat{\boldsymbol{x}},t)\,, \quad \boldsymbol{F}(\hat{\boldsymbol{x}},t) = \boldsymbol{G}^{-1}(\boldsymbol{x},t) \tag{1.6}$$

where $\boldsymbol{F}$ is a Lagrangian tensor and $\mathbf{G}$ is an Eulerian tensor. The velocity and the acceleration term are now presented both for Lagrangian and Eulerian formulations. If we consider the the Lagrangian formulation we have

$$\mathbf{v}(\hat{\boldsymbol{x}},t) = \frac{\partial \boldsymbol{x}}{\partial t}(\hat{\boldsymbol{x}},t)\,, \qquad \mathbf{a}(\hat{\boldsymbol{x}},t) = \frac{\partial^2 \boldsymbol{x}}{\partial t^2}(\hat{\boldsymbol{x}},t)\,. \tag{1.7}$$

At the same time, if we consider the Eulerian formulation we have

$$\mathbf{v}(\boldsymbol{x},t) = \frac{\partial \boldsymbol{x}}{\partial t}(\mathcal{X}(\boldsymbol{x}),t)\,, \qquad \mathbf{a}(\boldsymbol{x},t) = \frac{\partial^2 \boldsymbol{x}}{\partial t^2}(\mathcal{X}(\boldsymbol{x}),t)\,. \tag{1.8}$$

In order to define properly the introduced properties, the mapping $\mathcal{X}(\hat{\boldsymbol{x}},t)$ has to be invertible and twice continuously differentiable, in addition to the properties shown in Definition 1.1. In order to guarantee the local invertibility of the mapping $\mathcal{X}$, the tensor $\boldsymbol{F}$ has to be non-singular, which means that $J = \det \boldsymbol{F} \neq 0$. In the physical applications of this model, we always consider mapping that preserves the orientation of the transformed quantity. This means that $J > 0$. In addition, if we are interested in volume-preserving transformations, we have $J = 1$. This is a characteristic of the incompressible structures. More generally, J represents the relation between the reference and the transformed volume. In particular, if we consider the infinitesimal volume $d\hat{\boldsymbol{V}}$ in the reference configuration and the one $d\boldsymbol{V}$ in the current configuration, we have

$$d\boldsymbol{V} = J\,d\hat{\boldsymbol{V}}\,. \tag{1.9}$$

It is now interesting to consider the vector transformation through $\mathcal{X}$. We consider a generic vector $d\hat{\boldsymbol{x}} \in \hat{\Omega}$, it can be demonstrated that the following properties hold

$$d\boldsymbol{x} = \boldsymbol{F}\,d\hat{\boldsymbol{x}}\,, \quad \|\,d\boldsymbol{x}\,\| = \sqrt{d\hat{\boldsymbol{x}}^T \boldsymbol{F}^T \boldsymbol{F} d\hat{\boldsymbol{x}}}\,, \tag{1.10}$$

with the difference between the two vectors that can be written as

$$\|\,d\boldsymbol{x}\,\|^2 - \|\,d\hat{\boldsymbol{x}}\,\|^2 = d\hat{\boldsymbol{x}}^T\,(\boldsymbol{F}^T \boldsymbol{F} - \boldsymbol{I})\,d\hat{\boldsymbol{x}} = d\hat{\boldsymbol{x}}^T 2\boldsymbol{E}\,d\hat{\boldsymbol{x}}\,, \tag{1.11}$$

where $\boldsymbol{E}$ is the *Green-Lagrange tensor*, that can be written as

$$\boldsymbol{E} = \frac{1}{2}(\boldsymbol{F}^T\boldsymbol{F} - \boldsymbol{I}) = \frac{1}{2}\big(\nabla\boldsymbol{\eta} + \nabla\boldsymbol{\eta}^T + \nabla\boldsymbol{\eta}^T\nabla\boldsymbol{\eta}\big)\,. \tag{1.12}$$

If we consider only small deformations the term $\nabla\boldsymbol{\eta}^T\nabla\boldsymbol{\eta}$ is negligible, since it involves the product between two small quantities. So we can introduce the *small deformation tensor* $\boldsymbol{D}$ as

$$\boldsymbol{D} = \frac{1}{2}\big(\nabla\boldsymbol{\eta} + \nabla\boldsymbol{\eta}^T\big)\,. \tag{1.13}$$

If we analyze the eigenvalues of the matrix $\boldsymbol{D}$, we find that the $det(\boldsymbol{D}-\lambda\boldsymbol{I})$ can be written as a polynomial, function of λ

$$p(\lambda) = det(\boldsymbol{D} - \lambda\boldsymbol{I}) = \lambda^3 + I_D\lambda^2 - II_D\lambda + III_D\,. \tag{1.14}$$

It can be demonstrated that the three coefficients of the polynomial I_D, II_D, III_D don't change with rotations of the coordinate system. In fact, they are called *principal invariants* of tensors. In particular, if we find the three eigenvalues λ_1, λ_2, λ_3 of $\boldsymbol{D}$ (posing $p(\lambda) = 0$), we can define the principle invariants as

$$
\begin{aligned}
I_D &= tr(\boldsymbol{D}) = \lambda_1 + \lambda_2 + \lambda_3 \\
II_D &= \frac{1}{2}\big(tr(\boldsymbol{D})\big)^2 - tr(\boldsymbol{D}^2) = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3 \\
III_D &= det(\boldsymbol{D}) = \lambda_1\lambda_2\lambda_3
\end{aligned}
\tag{1.15}
$$

In the following paragraphs, we will compute forces acting on surfaces in both reference and current coordinate systems, so we need some relations useful for our scope [6]. The relation between infinitesimal surface area elements in the current and reference configurations is known as Nanson formula

$$\boldsymbol{n}ds = J\boldsymbol{F}^{-T}\boldsymbol{N}dS\,, \tag{1.16}$$

where we defined $\boldsymbol{N}dS$ and $\boldsymbol{n}ds$ as the unit normal vector multiplied by the area element in the reference and deformed configuration, respectively. Note that $\boldsymbol{F}^{-T} = (\boldsymbol{F}^{-1})^T$. We can now use (1.16) to obtain the so-called *Piola identity*. To this aim, we use the divergence theorem

$$\int_{\hat{\Omega}_t} \nabla \cdot \psi dv = \int_{\hat{\Gamma}_t} \psi \cdot \boldsymbol{n}ds\,, \tag{1.17}$$

of an identically constant function $\psi = 1$, so that the volume integral vanishes, obtaining

$$\int_{\hat{\Gamma}_t} \boldsymbol{n} ds = \int_{\hat{\Gamma}_t} J\boldsymbol{F}^{-T}\boldsymbol{N} dS = \int_{\hat{\Omega}_t} \nabla \cdot (J\boldsymbol{F}^{-T}) dv = 0 \,, \qquad (1.18)$$

which gives the expression of the Piola identity

$$\nabla \cdot (J\boldsymbol{F}^{-T}) = 0 \,. \qquad (1.19)$$

Stress tensors

In this paragraph, we introduce some stress measures for both Eulerian and Lagrangian formalism. In particular, in the Eulerian description the Cauchy stress tensor is the physical measure of the stress per unit area of the deformed configuration most commonly used. However, some other stress measures must be introduced in order to describe continuum mechanics in the Lagrangian formalism. For this purpose, we consider the Eulerian area element ds. Let $\boldsymbol{n}$ be the external normal unit vector to the surface delimiting the portion of the body. The surface forces $d\boldsymbol{f}_s$ applied to the considered Eulerian area element can be written as

$$d\boldsymbol{f}_s = \boldsymbol{\sigma}\boldsymbol{n} ds \,, \qquad (1.20)$$

where $\boldsymbol{\sigma}$ is the Cauchy stress tensor. In a similar fashion, we introduce a stress tensor $\boldsymbol{P}$, called the first Piola-Kirchhoff stress tensor such that $d\boldsymbol{f}_s = \boldsymbol{P}\boldsymbol{N} dS$, that relates forces in the current configuration to areas in the reference configuration and in general is not symmetric. By using the Nanson formula (1.16) we obtain

$$d\boldsymbol{f}_s = \boldsymbol{\sigma} J\boldsymbol{F}^{-T}\boldsymbol{N} dS \implies \boldsymbol{P} = J\boldsymbol{\sigma}\boldsymbol{F}^{-T} \,. \qquad (1.21)$$

We now introduce the transformation used to obtain the first Piola-Kirchhoff stress tensor from the Cauchy stress tensor, which is called Piola transformation.

Let $\hat{f} : \hat{\Omega} \to \mathbb{R}$ be a regular function defined over the domain $\hat{\Omega}$ and $f : \Omega \to \mathbb{R}$ the image of $\hat{f}$ such that $\hat{f}(\hat{\boldsymbol{x}}) = \hat{f}(\mathcal{X}(\hat{\boldsymbol{x}}))$. So we obtain

$$\nabla \hat{f} = \boldsymbol{F}^T \nabla f \,. \qquad (1.22)$$

In the following, the divergence terms are taken with respect to the corresponding domain configuration unless stated otherwise. Using (1.22) we can now introduce the so called Piola tensor transformation.

Definition 1.3. *Let $\boldsymbol{\tau} : \Omega \to \mathbb{R}^{3\times3}$ be a generic tensor defined over the deformed domain Ω. The Piola transformation of $\boldsymbol{\tau}$ associated to the transformation $\mathcal{X}$ is the second order tensor $\hat{\boldsymbol{\Pi}} : \hat{\Omega} \to \mathcal{R}^{3\times3}$ defined as*

$$\hat{\boldsymbol{\Pi}}(\hat{\boldsymbol{x}}) = J(\hat{\boldsymbol{x}})\,\boldsymbol{\tau}\big(\mathcal{X}(\hat{\boldsymbol{x}})\big)\big[\boldsymbol{F}^T(\hat{\boldsymbol{x}})\big]^{-1} = J\boldsymbol{\tau}\boldsymbol{F}^{-T}\,, \qquad \forall \hat{\boldsymbol{x}} \in \hat{\Omega}\,. \qquad (1.23)$$

The main interest of the Piola transform is that it yields a simple relation between the divergences of the tensors $\boldsymbol{\tau}$ and $\hat{\boldsymbol{\Pi}}$. Now we introduce a theorem on the properties of the Piola transform, following the mathematical formulation in [32].

Theorem 1.1. *Let $\boldsymbol{\tau} : \Omega \to \mathbb{R}^{3\times3}$ and $\hat{\boldsymbol{\Pi}} : \hat{\Omega} \to \mathcal{R}^{3\times3}$ as in definition 1.3. Then, by recalling $J = \det(\boldsymbol{F}) = \det(\nabla\mathcal{X}(\hat{\boldsymbol{x}},t))$ we state that*

$$\nabla \cdot \hat{\boldsymbol{\Pi}}(\hat{\boldsymbol{x}}) = J\nabla \cdot \boldsymbol{\tau}\big(\mathcal{X}(\hat{\boldsymbol{x}})\big)\,, \qquad \forall \hat{\boldsymbol{x}} \in \hat{\Omega}\,, \qquad (1.24)$$

$$\hat{\boldsymbol{\Pi}}(\hat{\boldsymbol{x}})\hat{\boldsymbol{n}}dS = \boldsymbol{\tau}\big(\mathcal{X}(\hat{\boldsymbol{x}})\big)\boldsymbol{n}ds\,, \qquad \forall \hat{\boldsymbol{x}} \in \partial\hat{\Omega}\,. \qquad (1.25)$$

Moreover, the area element ds and dS at the points $\hat{\boldsymbol{x}} \in \partial\hat{\Omega}$ and $\boldsymbol{x} = \mathcal{X}(\hat{\boldsymbol{x}}) \in \partial\Omega$, with unit outer normal vectors $\hat{\boldsymbol{n}}$ and $\boldsymbol{n}$, respectively, are related by

$$\det \nabla\mathcal{X}(\hat{\boldsymbol{x}})\,\big|\nabla\mathcal{X}(\hat{\boldsymbol{x}})^{-T}\,\hat{\boldsymbol{n}}\big|\,dS = \big|Cof\,\nabla\mathcal{X}(\hat{\boldsymbol{x}})\,\hat{\boldsymbol{n}}\big|dS = ds\,, \qquad (1.26)$$

where $Cof\,(A) = (\det A)A^{-T}$.

Proof. We recall that $J = \det(\nabla\mathcal{X}(\hat{\boldsymbol{x}}))$, $\boldsymbol{F} = \nabla\mathcal{X}(\hat{\boldsymbol{x}},t)$, and the Piola identity $\nabla \cdot (J\boldsymbol{F}^{-T}) = 0$. Then the relations

$$\hat{\boldsymbol{\Pi}}_{ij}(\hat{\boldsymbol{x}}) = J\boldsymbol{\tau}_{ik}\big(\mathcal{X}(\hat{\boldsymbol{x}})\big)\big(\boldsymbol{F}(\hat{x})^{-T}\big)_{kj}\,,$$

imply that

$$\hat{\partial}_j\hat{\boldsymbol{\Pi}}_{ij}(\hat{\boldsymbol{x}}) = J\hat{\partial}_j\boldsymbol{\tau}_{ik}\big(\mathcal{X}(\hat{\boldsymbol{x}})\big)\big(\boldsymbol{F}(\hat{x})^{-T}\big)_{kj}\,,$$

since the other term vanishes as a consequence of the Piola identity. The "hat" notation $\hat{\ }$ applied to the operators denotes that the operator is applied to the reference configuration. Next, by the chain rule,

$$\hat{\partial}_j\boldsymbol{\tau}_{ik}\big(\mathcal{X}(\hat{\boldsymbol{x}})\big) = \partial_l\boldsymbol{\tau}_{ik}\big(\mathcal{X}(\hat{\boldsymbol{x}})\big)\big(\boldsymbol{F}(\hat{\boldsymbol{x}})\big)_{lj}\,,$$

and the relation between $\hat{\nabla} \cdot \hat{\boldsymbol{\Pi}}_{ij}(\hat{\boldsymbol{x}})$ and $\nabla \cdot \boldsymbol{\tau}_{ik}\big(\mathcal{X}(\hat{\boldsymbol{x}})\big)$ follows from

$$(\boldsymbol{F}(\hat{\boldsymbol{x}}))_{li}(\boldsymbol{F}(\hat{\boldsymbol{x}})^{-T})_{ki} = \delta_{ik}\,,$$

where δ_{ik} denotes the Kronecker delta. Next, considering the relation $d\boldsymbol{x} = J d\hat{\boldsymbol{x}}$, along with the divergence theorem for tensor fields expressed over arbitrary $A \subseteq \hat{\Omega}$ and the formula for changes of variables in multiple integrals, we obtain

$$\int_{\partial A} \hat{\boldsymbol{\Pi}}(\hat{\boldsymbol{x}}) \hat{\boldsymbol{n}} \, d\hat{a} = \int_A \hat{\nabla} \cdot \hat{\boldsymbol{\Pi}}(\hat{\boldsymbol{x}}) \, d\hat{\boldsymbol{x}} = \int_A \nabla \cdot \boldsymbol{\tau}\big(\mathcal{X}(\hat{\boldsymbol{x}})\big) \, J \, d\boldsymbol{x}$$
$$= \int_{\mathcal{X}(A)} \nabla \cdot \boldsymbol{\tau}\big(\mathcal{X}(\hat{\boldsymbol{x}})\big) \, d\boldsymbol{x} = \int_{\mathcal{X}(\partial A)} \boldsymbol{\tau}\big(\mathcal{X}(\hat{\boldsymbol{x}})\big) \boldsymbol{n} \, da \, .$$

Since the domain A is arbitrary, we can prove the relation $\hat{\boldsymbol{\Pi}}(\hat{\boldsymbol{x}}) \hat{\boldsymbol{n}} \, d\hat{a} = \boldsymbol{\tau}\big(\mathcal{X}(\hat{\boldsymbol{x}})\big) \boldsymbol{n} \, da$. As a special case, we can also obtain the relation

$$J \boldsymbol{F}^{-T} \hat{\boldsymbol{n}} \, d\hat{a} = \boldsymbol{n} \, da \, ,$$

between the area elements $d\hat{a}$ and da by taking the Piola transform of the unit tensor $\boldsymbol{I}$. Thus, since $|\boldsymbol{n}| = 1$, da can be seen as the Euclidean norm of $J \boldsymbol{F}^{-T} \hat{\boldsymbol{n}} \, d\hat{a}$. This demonstrates the relation (1.26) by considering $dS = d\hat{a}$ and $ds = da$ as

$$ds = J |\boldsymbol{F}^{-T} \hat{\boldsymbol{n}}| dS = \det F |\boldsymbol{F}^{-T} \hat{\boldsymbol{n}}| dS$$

$\square$

In addition to Definition 1.3 and Theorem 1.1, if we consider the Cauchy stress tensor $\boldsymbol{\tau} = \boldsymbol{\sigma}$, we obtain

$$\hat{\boldsymbol{\Pi}}(\hat{\boldsymbol{x}}) = J \boldsymbol{\sigma} \boldsymbol{F}^{-T} \, . \tag{1.27}$$

In this context, we can derive the relationship between the divergence of a field taken in the two configurations, using the Piola identity (1.19) and the chain rule, as

$$\nabla \cdot \hat{\boldsymbol{\Pi}} = J \nabla \cdot \boldsymbol{\sigma} \, . \tag{1.28}$$

We reecall that the divergence terms are taken with respect to the corresponding domain configuration.

Lastly, we introduce the second Piola-Kirchhoff stress tensor $\boldsymbol{S}$ defined as

$$\boldsymbol{S} = \boldsymbol{F}^{-1} \boldsymbol{P} = \boldsymbol{F}^{-1} J \boldsymbol{\sigma} \boldsymbol{F}^{-T} \, . \tag{1.29}$$

All the introduced relations are useful to transform integrals involving the divergence of a tensor from the current configuration to the reference one and vice versa. For this reason all the presented tensors and transformation are essential for the FSI modeling.

The ALE formulation

In the FSI problems both the Lagrangian and the Eulerian formulations are not well suited for the numerical simulation of the fluid. In particular, the Eulerian approach contrasts with our needs of having a fluid mesh that moves following the deformation of the fluid-solid interface. However, some works on fluid-structure interaction problem are based on an Eulerian description of the fluid, with the transformation embedded in the simulated equations. In this work, this approach is not used, therefore an intermediate and more general description that combines the advantages of the Eulerian and Lagrangian approaches is introduced. So we introduce a computational moving domain $\omega(t)$, constrained to follow the interface fluid-solid on the boundary $\partial\omega(t)$. The name of this formulation is *Arbitrary Lagrangian Eulerian* (ALE) [33, 34].

Definition 1.4. *Let $\hat{\omega} = \omega(t = 0)$ be the reference domain for ω, we define the ALE map as the transformation*

$$\mathcal{A} : \hat{\omega} \times \mathbb{R}^+ \to \mathbb{R}^3 \,, \quad (\hat{\boldsymbol{x}}, t) \mapsto \hat{\boldsymbol{x}} = \mathcal{A}(\hat{\boldsymbol{x}}, t) \,, \tag{1.30}$$

where $\hat{\omega} \subset \mathbb{R}^3$ and such that $\hat{\omega}(t) = \mathcal{A}(\hat{\omega}, t)$.

We can now define the *computational domain velocity*, or *ALE velocity*, the vector

$$\hat{\boldsymbol{w}}(\hat{\boldsymbol{x}}, t) = \frac{\partial \mathcal{A}}{\partial t}(\hat{\boldsymbol{x}}, t) \,, \quad \forall \hat{\boldsymbol{x}} \in \hat{\omega} \,. \tag{1.31}$$

With this notation, we can now map every point of the current domain ω into the reference one $\hat{\omega}$. Furthermore the Lagrangian and the Eulerian cases can be seen as a particular case of the ALE description. In fact, in the Eulerian formulation the computational domain is fixed ($\omega(t) = \hat{\omega}(0)$) and the ALE velocity is null $\hat{\boldsymbol{w}} = 0$. Vice versa, in the Lagrangian formulation the computational domain moves with the same velocity of the material domain and the ALE and mesh velocities coincide.

Let now analyze the time derivative of a generic field $q(\boldsymbol{x}, t)$ for all the cited notations (Eulerian, Lagrangian, and ALE). If q is an Eulerian field, we define the *Eulerian derivative* as

$$\frac{\partial q}{\partial t}(\boldsymbol{x}, t) \,, \quad \boldsymbol{x} \in \Omega(t) \,. \tag{1.32}$$

Therefore only the temporal variation of q in the spatial fixed coordinate $\boldsymbol{x}$ is evaluated. This definition coincides with the standard definition of partial

derivative. If now we consider $\hat{q}$ as the Lagrangian description of q, we can define the *Lagrangian derivative* of q as

$$\frac{Dq}{Dt}(\cdot,t) = \frac{\partial \hat{q}}{\partial t}(\cdot,t) \circ \hat{\boldsymbol{\varphi}}_t^{-1} = \frac{d}{dt}q(\hat{\boldsymbol{\varphi}}(\hat{\boldsymbol{x}},t),t)\,, \quad \text{with } \boldsymbol{x} = \hat{\boldsymbol{\varphi}}(\hat{\boldsymbol{x}},t)\,. \quad (1.33)$$

In this case the variation is a function of the position $\boldsymbol{x}$ at the time t. By applying the chain rule for derivation we obtain

$$\frac{Dq}{Dt}(\hat{\boldsymbol{x}},t) = \frac{\partial q}{\partial t} + \frac{\partial q}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{x}}{\partial t} = \frac{\partial q}{\partial t} + \boldsymbol{v} \cdot \nabla q\,. \quad (1.34)$$

The first contribution refers to the time dependence of q, the Eulerian derivative (1.32) while the second one takes into account the advection of the field q due to the material velocity $\boldsymbol{v}$. Lastly, the ALE time derivative of a Eulerian field reads

$$\left.\frac{\partial q}{\partial t}\right|_{\mathcal{A}} = \frac{d}{dt}\big(q(\mathcal{A}(\hat{x},t),t)\big) = \frac{\partial q}{\partial t} + \boldsymbol{w} \cdot \nabla q\,, \quad (1.35)$$

where the first term is the Eulerian derivative and $\boldsymbol{w}$ is the Eulerian representation of the computational domain velocity. In this case, the second contribution takes into account the advection of the field q due to the motion of the mesh.

1.1.2   The conservation laws

All the equations of the FSI problem are partial differential equations derived from the conservation laws. In order to describe the behavior of the fluid and the solid, we need to impose the balance of mass, momentum and energy (in the case of non-isothermal problems). In this section, a brief introduction to the conservation laws is reported.

Let now consider a generic extensive quantity ψ defined over the generic domain $\Omega_t = \{\boldsymbol{x} : \boldsymbol{x} = \mathcal{X}(\hat{\boldsymbol{x}},t)\,, \hat{\boldsymbol{x}} \in \Omega\,\}$. The general conservation equation for ψ can be written as

$$\begin{aligned}
\frac{d}{dt}\int_{\Omega_t} \rho\psi d\boldsymbol{x} = &-\int_{\partial\Omega_t} \boldsymbol{J}(\psi) \cdot \hat{\boldsymbol{n}}dA + \int_{\Omega_t} \rho S(\phi,\boldsymbol{x})d\boldsymbol{x} \\
&+ \int_{\Omega_t}\int_{\Omega_t} K(\psi)(\boldsymbol{x}',\boldsymbol{x})\psi(\boldsymbol{x}',t)d\boldsymbol{x}'d\boldsymbol{x}\,,
\end{aligned} \quad (1.36)$$

where $\partial\Omega_t$ is the external boundary of Ω_t, $\boldsymbol{J}(\psi)$ is the current of the variable ψ through $\partial\Omega_t$, and $S(\phi,\boldsymbol{x})$ is a source term. Therefore, the global conservation

of ψ is determined by the flux density $\boldsymbol{J}_\psi$, the external source $\rho S(\psi, \boldsymbol{x})$ and the internal source $\int_{\Omega_t} K(\psi)(\boldsymbol{x}', \boldsymbol{x})\psi(\boldsymbol{x}', t)d\boldsymbol{x}'$, as can be seen in (1.36).

Now we introduce the *Reynolds transport theorem*, following the demonstration in [35] for moving boundaries, in order to state the balance laws. In the following, the Euclidean space will be denoted by $\mathcal{E}$. A moving volume region with boundary in $\mathcal{E}$ is a family of subsets $\{\mathcal{D}(t) \,|\, t \in T\}$ in $\mathcal{E}$, where each set $\mathcal{D}(t)$ is a regular region with a boundary $\partial\mathcal{D}(t)$ such that $\{\partial\mathcal{D}(t) \,|\, t \in T\}$ is a closed moving surface region in $\mathcal{E}$ with an orientation $\boldsymbol{n}(\cdot, t)$ directed out of the region $\mathcal{D}(t)$. Obviously, following the notation in the previous part of this chapter, we can set $\mathcal{D}(t) = \Omega_t(t)$.

Theorem 1.2. *Suppose $\{\Omega_t(t) \,|\, t \in T\}$, where T is an open interval of $\mathbb{R}$, is a moving region with boundary in $\mathcal{E}$. Let $\psi(\cdot, t) : \Omega_t(t) \to \mathbb{R}$ be a scalar field on $\Omega_t(t)$. Then*

$$\frac{d}{dt}\int_{\Omega_t} \psi(\boldsymbol{x}, t)\, d\boldsymbol{v}(\boldsymbol{x}) = \int_{\Omega_t} \frac{\partial\psi}{\partial t}\, d\boldsymbol{v}(\boldsymbol{x}) + \int_{\partial\Omega_t} \psi(\boldsymbol{x}, t)\boldsymbol{u}_n(\boldsymbol{x}, t)\, da(x)\,, \quad (1.37)$$

where $\boldsymbol{u}_n$ is the normal speed of $\partial\Omega_t$.

Proof. Consider $\tau \in [t_0, t]$ and the differential form $\omega(\boldsymbol{x}, t) = \psi(\boldsymbol{x}, \tau)d\boldsymbol{v}(\boldsymbol{x})$, where $d\boldsymbol{v}(\boldsymbol{x})$ is the volume form on $\mathcal{E}$. The Stokes formula gives [36]

$$\int_{\Omega_t} d\omega(\boldsymbol{x}, \tau) = \int_{\partial\Omega_t} \omega(\boldsymbol{x}, \tau)\,. \quad (1.38)$$

Let now consider the translation vector space $\mathcal{V}$, with a fixed orthonormal base $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$. Then, we define $\hat{\psi}(x, y, z, \tau) = \psi(x_0 + \boldsymbol{a}_1 x + \boldsymbol{a}_2 y + \boldsymbol{a}_3 z, \tau)$, where x_0 is some fixed point in $\mathcal{E}$. Since $d(d\boldsymbol{v}(\boldsymbol{x})) = 0$, we have

$$d\omega(\boldsymbol{x}, \tau) = d\psi(\boldsymbol{x}, \tau)d\psi(\boldsymbol{x}, \tau) \wedge d\boldsymbol{v}(\boldsymbol{x})$$
$$= \left(\frac{\partial\hat{\psi}}{\partial x}dx + \frac{\partial\hat{\psi}}{\partial y}dy + \frac{\partial\hat{\psi}}{\partial z}dz + \frac{\partial\hat{\psi}}{\partial\tau}d\tau\right) \wedge (dx \wedge dy \wedge dz)$$
$$= -\frac{\partial\hat{\psi}}{\partial\tau}d\tau \wedge dx \wedge dy \wedge dz = -\frac{\partial\psi}{\partial\tau}d\boldsymbol{v}(\boldsymbol{x}) \wedge d\tau\,,$$

where the symbol $\wedge$ accounts the derivative of the product terms. Then, we have

$$\int_{\Omega_t(t)} d\omega(\boldsymbol{x}, \tau) = \int_{\Omega_t(t)} -\frac{\partial\psi}{\partial\tau}d\boldsymbol{v}(\boldsymbol{x}) \wedge d\tau = -\int_{t_0}^{t}\left(\int_{\Omega_t(t)} \frac{\partial\psi}{\partial\tau}d\boldsymbol{v}(\boldsymbol{x}) \wedge d\tau\right)d\tau\,.$$
$$(1.39)$$

Moreover

$$\int_{\partial\Omega_t} d\omega(\boldsymbol{x},\tau) = \int_{\Omega_t(t_0)} \psi(\boldsymbol{x},t_0)d\boldsymbol{v}(\boldsymbol{x}) + \int_{\mathcal{T}(t)} \psi(\boldsymbol{x},\tau)d\boldsymbol{v}(\boldsymbol{x})$$
$$- \int_{\Omega_t(t)} \psi(\boldsymbol{x},t)d\boldsymbol{v}(\boldsymbol{x})\,,$$

where $\mathcal{T}(t) = \{(\boldsymbol{x},\tau)|\boldsymbol{x} \in \Omega_t(\tau), \tau \in [t_0,t]$. By considering a local coordinate system for the surface $\partial\Omega_t(t)$, it can be shown [35] that

$$d\boldsymbol{v}(\boldsymbol{x}) = dx \wedge dy \wedge dz = \boldsymbol{u}_n da(x) \wedge d\tau\,,$$

where $\boldsymbol{u}_n = \boldsymbol{u} \cdot \boldsymbol{n}$ is the normal speed on $\partial\Omega_t$, $\forall t \in [t_0,t]$. Thus

$$\int_{\mathcal{T}(t)} \Big(\psi(\boldsymbol{x},\tau)d\boldsymbol{v}(\boldsymbol{x})\Big)d\tau = \int_{\mathcal{T}(t)} \psi(\boldsymbol{x},\tau)\boldsymbol{u}_n(\boldsymbol{x},t)\,da(\boldsymbol{x}) \wedge d\tau$$
$$= \int_{t_0}^t \Big(\int_{\partial\Omega_t(\tau)} \psi(\boldsymbol{x},\tau)\boldsymbol{u}_n(\boldsymbol{x},\tau)\,da(\boldsymbol{x})\Big)d\tau\,, \tag{1.40}$$

and then by combining (1.38) - (1.40) we obtain

$$-\int_{t_0}^t \Big(\int_{\Omega_t()t} \frac{\partial\psi(\boldsymbol{x},\tau)}{\partial\tau}d\boldsymbol{v}(\boldsymbol{x})\Big)d\tau = \int_{\Omega_t(t_0)} \psi(\boldsymbol{x},t_0)d\boldsymbol{v}(\boldsymbol{x})$$
$$+ \int_{t_0}^t \Big(\int_{\partial\Omega_t(\tau)} \psi(\boldsymbol{x},\tau)\boldsymbol{u}_n(\boldsymbol{x},t)\,da(\boldsymbol{x})\Big)d\tau - \int_{\Omega(t)} \psi(\boldsymbol{x},t)d\boldsymbol{v}(\boldsymbol{x})\,. \tag{1.41}$$

By differentiating the relation (1.41) with respect to t we obtain (1.37). $\square$

Another commonly used formulation for the Reynolds theorem is

$$\frac{d}{dt}\int_{\Omega_t} \psi(\boldsymbol{x},t)d\boldsymbol{x} = \int_{\Omega_t} \Big(\frac{\partial\psi(\boldsymbol{x},t)}{\partial t} + \nabla \cdot (\psi(\boldsymbol{x},t)\boldsymbol{v})\Big)d\boldsymbol{x}\,. \tag{1.42}$$

Mass conservation

Firstly, we consider the conservation of the mass. Let consider the mass m contained in a fixed region $V(t) \subset \mathbb{R}^3$, then

$$m = \int_V \rho(\boldsymbol{x},t)\,d\boldsymbol{x}\,, \tag{1.43}$$

where $\rho(\boldsymbol{x},t)$ is the mass density in the Eulerian description. Then, by applying (1.42) we obtain

$$\frac{dm}{dt} = \frac{d}{dt}\int_V \rho(\boldsymbol{x},t)\,d\boldsymbol{x} = \int_V \Big[\frac{\partial\rho}{\partial t} + \nabla \cdot (\rho\boldsymbol{u})\Big]d\boldsymbol{x} = 0\,. \tag{1.44}$$

The equation (1.44) is the integral form of the mass conservation equation in Eulerian form. Let now consider small enough involved fields. Then, we can write the local form

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{u}) \,,$$

and, in the case of constant ρ we obtain the incompressibility constraint

$$\nabla \cdot \boldsymbol{u} = 0 \,.$$

All the presented equations refer to the Eulerian description. Now, we introduce the corresponding form of the mass conservation in the Lagrangian formulation. For this purpose, we introduce a fixed set of particles V^*. The mass conservation turns to

$$\frac{dm}{dt} = \frac{d}{dt} \int_{V^*} \rho \, d\boldsymbol{x}^* = \int_{V^*} \frac{\partial \rho J}{\partial t} \, d\boldsymbol{x} = 0 \,.$$

Since V can be chosen arbitrarily, the results to any material point can be localized as

$$\frac{\partial}{\partial t}(\rho J) = 0 \,, \qquad \forall t \geq 0 \,.$$

Therefore, the product ρJ is only a function of the considered material point, and does not change in time. If we consider an FSI simulation, then we can state that at $t = 0$ the structure is not deformed, then $J = 1$. Let $\rho_0 = \rho_0(\boldsymbol{X})$ be the structural mass density in the undeformed configuration as a function of the material point $\boldsymbol{X}$, then we recover the following expression of the continuity equation in the Lagrangian form

$$\rho_0 = \rho J \,,$$

then by knowing the displacement field of the structure (i.e. J is known) it is possible to obtain the density at a material point in the current configuration.

Momentum conservation

The principle of conservation of linear momentum (or Newton's Second Law of motion) states that the time rate of change of linear momentum of a given set of particles is equal to the vector sum of all the external forces acting on the particles of the set, provided Newton's Third Law of action and reaction governs the internal forces [37]. Therefore, by applying Newton's Second Law to the to a material domain $V(t)$ we obtain the conservation of

momentum equation. Such an equation states that the resultant of external forces (volume forces $\boldsymbol{F}_v$ and surface forces F_s) is equal to the rate of change of the total linear momentum of the system. The momentum equation can be written as

$$\frac{d}{dt} \int_{(t)} \rho \boldsymbol{u} \, d\boldsymbol{x} = \boldsymbol{F}_v + \boldsymbol{F}_s \,.$$

The volume forces can be obtained by integrating over the volume the specific force $\boldsymbol{f}_v$ multiplied by the material density

$$\boldsymbol{F}_v = \int_{V(t)} \rho \boldsymbol{f}_v \, d\boldsymbol{x} \,.$$

The surface forces can be obtained as the surface integral over the boundary $\partial V(t)$ of the Cauchy stress tensor $\boldsymbol{\sigma} : \Omega(t) \to \mathbb{R}^{3 \times 3}$

$$\boldsymbol{F}_s = \int_{\partial V(t)} \boldsymbol{\sigma} \boldsymbol{n} \, dS \,,$$

where $\boldsymbol{n}$ is the normal to the surface dS.

Now, by applying the Reynolds transport theorem and the divergence theorem to the surface forces, we obtain

$$\begin{aligned}
\frac{d}{dt} \int_{V(t)} \rho \boldsymbol{u} &= \int_{V(t)} \left[\frac{\partial \rho \boldsymbol{u}}{\partial t} + \nabla \cdot (\rho \boldsymbol{u} \boldsymbol{u}) \right] d\boldsymbol{x} \\
&= \int_{V(t)} \rho \boldsymbol{f}_v \, d\boldsymbol{x} + \int_{\partial V(t)} \boldsymbol{\sigma} \boldsymbol{n} \, dS \\
&= \int_{V(t)} (\rho \boldsymbol{f}_v + \nabla \cdot \boldsymbol{\sigma}) \, d\boldsymbol{x} \,.
\end{aligned}$$

Therefore

$$\int_{V(t)} \left[\frac{\partial \rho \boldsymbol{u}}{\partial t} + \nabla \cdot (\rho \boldsymbol{u} \boldsymbol{u}) \right] d\boldsymbol{x} = \int_{V(t)} (\rho \boldsymbol{f}_v + \nabla \cdot \boldsymbol{\sigma}) \, d\boldsymbol{x} \,.$$

Since the volume $V(t)$ is arbitrarily chosen we can recover the local formulation of the balance of momentum in Eulerian form

$$\rho \frac{\partial \boldsymbol{u}}{\partial t} + \rho (\boldsymbol{u} \cdot \nabla) \boldsymbol{u} = \rho \boldsymbol{f}_v + \nabla \cdot \boldsymbol{\sigma} \,, \tag{1.45}$$

in $\Omega(t)$ and for $t > 0$. Now, the Lagrangian formulation of (1.45) can be derived by mapping back the integrals on the undeformed configuration Ω, obtaining

$$\frac{d}{dt} \int_\Omega \rho J \boldsymbol{u} \, d\boldsymbol{x} = \int_\Omega \rho J \boldsymbol{f}_v \, d\boldsymbol{x} + \int_\Omega J \nabla \cdot \boldsymbol{\sigma} \, d\boldsymbol{x} \,.$$

Then, since Ω is an arbitrary volume, we obtain

$$\rho_0 \frac{\partial \boldsymbol{u}}{\partial t} = \rho_0 \boldsymbol{f}_v + J \nabla \cdot \boldsymbol{\sigma} \qquad \text{in } \Omega\,,$$

where the divergence is taken with respect to the deformed coordinates. It is possible to transform the term $J \nabla \cdot \boldsymbol{\sigma}$ into its Lagrangian form using (1.28), and obtaining

$$\rho_0 \frac{\partial \boldsymbol{u}}{\partial t} = \rho_0 \boldsymbol{f}_v + \nabla \cdot \boldsymbol{P} \qquad \text{in } \Omega\,,$$

where $\boldsymbol{P} = J \boldsymbol{\sigma} \boldsymbol{F}^{-T}$ is the first Piola-Kirchhoff stress tensor. In conclusion, we can write the momentum balance equation in the ALE formulation

$$\rho \frac{\partial \boldsymbol{u}}{\partial t}\big|_{\mathcal{A}} + \rho[(\boldsymbol{u} - \boldsymbol{w}) \cdot \nabla] = \rho \boldsymbol{f}_v + \nabla \cdot \boldsymbol{\sigma}\,, \qquad (1.46)$$

in $\Omega(t)$ and for $t > 0$. This formulation will be useful for the formulation of FSI problems.

1.2 Constitutive models for FSI

The main difference between a fluid and a solid is the macroscopic reaction to external forces. The conservation equations presented in the last section are undetermined. Then, it is necessary to bound such equations with constitutive models both for fluid and solid. In this section, some of the most used constitutive equations for fluids and solids are presented. In particular, in order to introduce properly the fluid-structure interaction simulations, we will link stress to strain and strain to kinematic variables. The models will be presented firstly in the strong formulation, then the weak formulation will be considered to couple the fluid and the solid equations into the FSI system.

1.2.1 Strong formulation for Newtonian fluids

We now focus on the fluid sub-problem and, in particular, we introduce a model for Newtonian fluids. This class of fluids is the simplest to solve, since the viscosity does not depend on the stress state and on the velocity $\boldsymbol{u}$. In particular, in the FSI framework, it is important to determine the relation between the Cauchy stress tensor $\boldsymbol{\sigma}$ and the fluid velocity. The Cauchy stress tensor can be decomposed in the hydrostatic and the deviatoric components. The hydrostatic component $\boldsymbol{\sigma}_{\text{hyd}} = -p_f \boldsymbol{I}$ tends to change the volume of the

stressed body. The deviatoric component is responsible for the fluid distortion rate of deformations $\boldsymbol{\sigma}_{\mathrm{dev}} = \boldsymbol{\tau}$. In this framework, the pressure p_f is Lagrange multiplier associated to the incompressibility constraint $\nabla \cdot \boldsymbol{u} = 0$. It can be written as

$$p_f = -\frac{\mathrm{tr}(\boldsymbol{\sigma})}{3} = -\frac{I_D}{3}\,,$$

where $\mathrm{tr}(\boldsymbol{\sigma})$ indicates the trace of the Cauchy tensor, and I_D the first invariant of the stress tensor introduced in (1.15). On the other hand, the deviatoric component is modeled as

$$\boldsymbol{\tau} = 2\mu_f \boldsymbol{D} - \frac{2}{3}\mu_f (\nabla \cdot \boldsymbol{u})\boldsymbol{I}\,, \qquad (1.47)$$

where μ_f is the fluid dynamic viscosity, and $\boldsymbol{D}$ is the fluid deformation tensor, that can be written as

$$\boldsymbol{D} = \frac{1}{2}\big(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T\big)\,.$$

If an incompressible fluid is considered, by using the incompressibility constraint $\nabla \cdot \boldsymbol{u} = 0$ the second term of (1.47) vanishes, obtaining $\boldsymbol{\tau} = 2\mu_f \boldsymbol{D}$.

Therefore, the Cauchy stress tensor for a Newtonian incompressible fluid reads

$$\boldsymbol{\sigma}_f = -p_f \boldsymbol{I} + \mu_f \big(\nabla \boldsymbol{u}_f + (\nabla \boldsymbol{u}_f)^T\big)\,. \qquad (1.48)$$

Note that for a Newtonian fluid, viscosity operates as diffusion of momentum. The presented model for Newtonian fluids is suitable for various common fluids, such as water and air, under ordinary conditions. However, some fluids do not follow Newton's law of viscosity, i.e., constant viscosity independent of stress. In this work, we will consider only Newtonian fluids for the implementation of the fluid-structure interaction models presented in the following sections.

1.2.2   Strong formulations for hyperelestic solid

In this section, the mathematical formulation for the modeling of hyperelestic solid structures is presented. However, since in the following a membrane model that behaves like an elastic material is presented, we present now a brief introduction to the elastic materials.

Elastic materials

An elastic material is a solid for which the stress field depends only on the current state of deformation and not on the deformation history. Let now consider a generic particle $\boldsymbol{X}$ (in Lagrangian formulation) and the deformation gradient $\boldsymbol{F}$, defined in (1.4). Therefore, under such conditions, any stress measure at $\boldsymbol{X}$ is a function of the gradient $\boldsymbol{F}$ associated with that particle. Now we recall the first Piola-Kirchhoff stress tensor $\boldsymbol{P}$ introduced above. We may express the constitutive equation as

$$\boldsymbol{P} = \boldsymbol{P}(\boldsymbol{F}(\boldsymbol{X}, t), \boldsymbol{X}) \,.$$

Moreover, if $\boldsymbol{P}$ and the mass density ρ_0 are independent of the particle position, the material is *homogeneous* and we have $\boldsymbol{P} = \boldsymbol{P}(\boldsymbol{F}(\boldsymbol{X}, t))$. Solid materials with the same response to deformation in all directions are called *isotropic*, and follow the relation

$$\boldsymbol{P}(\boldsymbol{F}\boldsymbol{Q}) = \boldsymbol{Q}^T \boldsymbol{P}(\boldsymbol{F})\boldsymbol{Q} \,,$$

where $\boldsymbol{Q}$ is the orthogonal matrix associated to the rigid body rotation.

Hyperelastic materials

We consider now the case where the work done by the stress during a deformation process is dependent only on the initial and final state. In this case the material is called *hyperelastic*, and its behavior is said to be path-independent. In this work the inelastic materials are not studied, however interested reader can see [38]. As a consequence of the described path-independent behavior, the *stored strain energy density function* per unit volume of the undeformed configuration Ψ can be established as the work done by the stress tensor from the initial position at time t_0 to the current position at time t as

$$\Psi(\boldsymbol{F}(\boldsymbol{X}), \boldsymbol{X}) = \int_{t_0}^{t} \boldsymbol{P}(\boldsymbol{F}(\boldsymbol{X}), \boldsymbol{X}) \,:\, \dot{\boldsymbol{F}} \, dt \,.$$

Note that different forms of $\Psi(\boldsymbol{F})$ lead to different constitutive relationships between stress and strain. Under these hypotheses, the first Piola-Kirchhoff stress tensor can be obtained by differentiating Ψ with respect to $\boldsymbol{F}$ as

$$\boldsymbol{P}(\boldsymbol{F}) = \frac{\partial \Psi(\boldsymbol{F}(\boldsymbol{X}), \boldsymbol{X})}{\partial \boldsymbol{F}} \,.$$

For a homogeneous, isotropic and frame indifference material, the potential energy depends only upon the deformation gradient $\boldsymbol{F}$. Besides we require for convenience that the strain-energy function vanishes in the reference configuration. However, physical observations show that the strain-energy function Ψ increases with deformation. Therefore we impose that $\Psi(\boldsymbol{F}) \geq 0$.

Now we recall the second Piola-Kirchhoff stress tensor $\boldsymbol{S}$ (see equation (1.29)). It can be obtained as a function of the right Cauchy-Green deformation tensor $\boldsymbol{C}$. We also recall that $\boldsymbol{S} = \boldsymbol{F}^{-1}\boldsymbol{P}$ and $\boldsymbol{C} = \frac{1}{2}(\boldsymbol{F}\boldsymbol{F}^{-T} - \boldsymbol{I})$. The tensor $\boldsymbol{S}$ can also be written as

$$\boldsymbol{S} = 2\frac{\partial\Psi(\boldsymbol{C}, \boldsymbol{X})}{\partial\boldsymbol{C}} = \frac{\partial\Psi(\boldsymbol{E}, \boldsymbol{X})}{\partial\boldsymbol{E}}$$

The Cauchy stress tensor can be computed from the equations of the first and second Piola-Kirchhoff tensors (equations (1.21) and (1.29), respectively).

We define now the *tensor of elastic moduli* $\mathcal{C}$ as the second derivative of Ψ with respect to $\boldsymbol{E}$

$$\mathcal{C}(\boldsymbol{E}) = \frac{\partial^2\Psi}{\partial\boldsymbol{E}\partial\boldsymbol{E}} .$$

The tensor $\mathcal{C}(\boldsymbol{E})$ is a four-rank tensor independent of the state of deformation for hyperelastic models. His number of entries can be reduced to only two by considering a symmetric and isotropic case.

We introduce now some of the most common quantities used to determine the properties of the materials. The *Young modulus E* is defined as the ratio between stress and strain in a material in the linear elastic regime. The *Poisson coefficient ν* is the opposite of the ratio between transverse strain and axial strain. Usually, $\nu \in [0, 0.5]$. We have $\nu \to 0.5$ for nearly incompressible materials. When $\nu \to 0$, the material shows null lateral expansion when compressed. The *Bulk modulus k* is the indicator of the resistance to compression of a solid and is defined as $k = -V\dfrac{\partial p}{\partial V}$, where p is pressure and V is the initial volume of the substance. We also introduce the *Lamé parameters* λ and μ, which are generally referred to as Lamé first parameter and Lamé second parameter, respectively. In homogeneous and isotropic materials, Hooke's law can be written as

$$\boldsymbol{\sigma} = 2\mu\boldsymbol{\varepsilon} + \lambda\mathrm{tr}(\boldsymbol{\varepsilon})\boldsymbol{I} ,$$

where $\boldsymbol{\varepsilon}$ is the strain tensor. The second Lamé parameter μ is also known as shear modulus G in the context of elasticity. Note that homogeneous isotropic linear elastic materials have their elastic properties uniquely determined by any two variables among E, ν, λ and μ.

The incompressible hyperelastic solid. We now consider the incompressible case for hyperelastic solid. We have

$$J = \det \boldsymbol{F} = III_C = 1\,.$$

Moreover, from the tensor calculus we have that the derivative of the determinant of the second-order tensor $\boldsymbol{F}$ is

$$\frac{\partial}{\partial \boldsymbol{F}} \det(\boldsymbol{F}) = \det(\boldsymbol{F})\boldsymbol{F}^{-T} = J\boldsymbol{F}^{-T}\,.$$

Thus, considering the introduced relations, along with

$$\Psi(\boldsymbol{F}(\boldsymbol{X}), \boldsymbol{X}) = \Psi(\boldsymbol{F}) - p(J - 1)\,,$$

we can write the first Piola-Kirchhoff tensor as

$$\boldsymbol{P} = \frac{\partial \Psi}{\partial \boldsymbol{F}} - pJ\boldsymbol{F}^{-T}\,,$$

where p is again the scalar hydrostatic pressure that can be seen as a Lagrangian multiplier that enforces the incompressibility constraint. Now, by recalling $\boldsymbol{S} = \boldsymbol{F}^{-1}\boldsymbol{P}$, we have [39]

$$\boldsymbol{S} = \boldsymbol{F}^{-1}\frac{\partial \Psi}{\partial \boldsymbol{F}} - pJ\boldsymbol{F}^{-1}\boldsymbol{F}^{-T} = 2\frac{\partial \Psi}{\partial \boldsymbol{C}} - \boldsymbol{F}^{-1}Jp\boldsymbol{F}^{-T}\,.$$

Thus, the Cauchy stress tensor can be written using the Piola transformation (1.23) with $\boldsymbol{\tau} = p\boldsymbol{I}$

$$\boldsymbol{\sigma} = -p\boldsymbol{I} + \boldsymbol{\sigma}^*\,.$$

In addition, the formulation can also be applied to the compressible case ($\nu_s \approx 0.4$)

$$\boldsymbol{\sigma} = -\lambda\Big(J - \frac{1}{J}\Big)\boldsymbol{I} + \frac{\boldsymbol{\sigma}^*}{J}\,. \tag{1.49}$$

The value of $\boldsymbol{\sigma}^*$ depends on the used model for the solid modeling. In the next paragraphs some models for the modeling of $\boldsymbol{\sigma}^*$ in hyperelastic solid applications are presented, namely the Saint Venant-Kirchhoff and the neo-Hookean model. In literature there are many other hyperelastic models, such as Mooney or Mooney-Rivlin materials. Interested reader can see [40].

Saint Venant-Kirchhoff model.   One of the simplest models for hyperelastic material is the Saint Venant–Kirchhoff model which is just an extension of the geometrically linear elastic material model to the geometrically nonlinear regime. Thus this model is characterized by a nonlinear strain energy function

$$\Psi(\boldsymbol{E}) = \frac{1}{2}\boldsymbol{E} : \mathcal{C}\boldsymbol{E}\,.$$

Therefore, the second Piola–Kirchhoff stress tensor $\boldsymbol{S}$ is linear with the respect of $\boldsymbol{E}$, namely $\boldsymbol{S}(\boldsymbol{E}) = \mathcal{C}\boldsymbol{E}$. If we consider an isotropic material, Ψ can be rewritten as

$$\Psi(\boldsymbol{F}) = \lambda(\mathrm{tr}(\boldsymbol{E}))^2 + 2\mu\,\mathrm{tr}(\boldsymbol{E}^2)\,,$$

Lemma 1.1. *The Cauchy stress tensor in the Saint Venant-Kirchhoff hypotheses can be written as*

$$\boldsymbol{\sigma} = \lambda(\nabla \cdot \boldsymbol{\eta})\boldsymbol{I} + \mu(\nabla\boldsymbol{\eta} + (\nabla\boldsymbol{\eta})^T)\,. \tag{1.50}$$

Proof. Under this hypothesis, the second Piola–Kirchhoff stress tensor can be written as

$$\boldsymbol{S}(\boldsymbol{E}) = \frac{\partial\Psi(\boldsymbol{E})}{\partial\boldsymbol{E}} = \lambda\,\mathrm{tr}(\boldsymbol{E})\boldsymbol{I} + 2\mu\boldsymbol{E}\,.$$

It is easy to adapt this model in order to deal with small deformations and incompressible materials. Let consider the small deformation tensor $\boldsymbol{D}$. We have

$$\mathrm{tr}(\boldsymbol{D}) = \nabla \cdot \boldsymbol{\eta}$$

Then, the restriction to small deformations of the Cauchy stress tensor can be linearized leading to (1.50). □

The equation (1.50) is also known as the Navier-Lameé equation. The St. Venant-Kirchhoff model is quite used in the computational dynamics of solids. It is easy to implement such a model, however, this model has some disadvantages, i.e the lack of any term preventing J to approach zero in the stored energy function. This model is therefore mainly adopted when dealing with small strains $\boldsymbol{E}$.

Neo-Hookean model.   We now introduce another model developed in order to enforce the Saint Venant-Kirchhoff one [41]. In fact, in contrast to linear elastic materials, the stress-strain curve of a neo-Hookean material is not linear. Instead, the relationship between applied stress and strain is initially linear, then at a certain stress level, the stress-strain curve presents a

plateau. For this reason, it is a good model to simulate plastics and rubber-like substances. For a Neo-Hookean material, with respect to the current configuration, the strain energy function takes the form

$$\Psi(I, J) = \frac{1}{2}\mu(J^{-2/3}I - 3) + \frac{1}{2}\Big(\lambda + \frac{2}{3}\mu\Big)\Big(\frac{1}{2}(J^2 - 1) - \ln(J)\Big). \quad (1.51)$$

We can now write the second Piola–Kirchhoff stress tensor as

$$\boldsymbol{S} = \mu J^{-2/3}\Big(\boldsymbol{I} - \frac{1}{3}I\boldsymbol{C}^{-1}\Big) + \frac{1}{2}\Big(\lambda + \frac{2}{3}\mu\Big)(J^2 - 1)\boldsymbol{C}^{-1}. \quad (1.52)$$

The equations (1.51) and (1.52) are written for a generic material. However, for incompressible materials we have $J = 1$.

Lemma 1.2. *The incompressible Cauchy stress tensor for neo-Hookean materials can be written as*

$$\boldsymbol{\sigma} = -p\boldsymbol{I} + \mu\boldsymbol{B} \quad (1.53)$$

where $\boldsymbol{B}$ is the left Cauchy-Green deformation tensor.

Proof. The strain energy in the incompressible case ($J = 1$) can be written as

$$\Psi(I, J) = \Psi(I) = \frac{1}{2}\mu(I - 3). \quad (1.54)$$

Then, the second Piola–Kirchhoff stress tensor can be written as

$$\boldsymbol{S} = \mu J^{-2/3}\Big(\boldsymbol{I} - \frac{1}{3}I\boldsymbol{C}^{-1}\Big) - p J\boldsymbol{C}^{-1}. \quad (1.55)$$

By considering (1.54) and (1.55) together with (1.28) one can easily obtain (1.52). $\qquad\square$

1.3 The fluid-structure interaction model

1.3.1 Basic notation

In order to keep the mathematical formulation of this work as consistent as possible, we introduce now some of the notation used in the rest of this thesis. The Sobolev spaces $W^{k,p}(\mathscr{C})$ are defined as the space of functions u on $\mathscr{C} \subset \mathbb{R}^n$ such that u and all its partial derivatives up to order k are in $L^p(\mathscr{C})$. For a given $s \in \mathbb{R}$, the standard Sobolev space of order s and $p = 2$ is denoted by $H^s(\mathscr{C})$. When $s \in \mathbb{Z}^+$, the inner product over $H^s(\mathscr{C})$ is denoted

by $(\cdot,\cdot)_s$ and $(\cdot,\cdot)$ denotes the inner product in $L^2(\mathscr{C})$. In the same way we define the natural norm as $\|\cdot\|_{s,\mathscr{C}}$. Whenever possible, the domain label in the norm is neglected. We use boldface notation for the corresponding Sobolev spaces of vector-valued functions, i.e. $\boldsymbol{H}^1(\Omega) = [H^1(\Omega)]^n$ is the space of the functions in $\mathbb{R}^n$ such that each component of $\boldsymbol{H}^1(\Omega)$ belongs to $H^1(\Omega)$. Dual spaces are denoted by $(\cdot)^*$. In particular, we consider the following spaces and subspaces

$$\boldsymbol{H}^1(\Omega) = \left\{ u_j \in L^2(\Omega) \,\Big|\, \frac{\partial u_j}{\partial x_k} \in L^2(\Omega) \quad \text{for } j,k = 1,\cdots,n \right\},$$

$$\boldsymbol{H}_0^1(\Omega) = \left\{ \boldsymbol{u} \in \boldsymbol{H}^1(\Omega) \,\Big|\, \boldsymbol{u} = 0 \quad \text{on } \partial\Omega \right\},$$

$$L_0^2(\Omega) = \left\{ p \in L^2(\Omega) \,\Big|\, \int_\Omega p^2 d\boldsymbol{x} = 0 \right\},$$

$$\boldsymbol{T}(\Omega) = \left\{ \boldsymbol{u} \in \boldsymbol{H}^1(\Omega) \,|\, \nabla \cdot \boldsymbol{u} = 0 \right\},$$

where $\partial\Omega$ represents the boundary of Ω.

Moreover, for a nonzero portion of the boundary $\Gamma_\alpha \subset \partial\Omega$, we introduce the subspace

$$\boldsymbol{H}_{\Gamma_\alpha}^1(\Omega) = \left\{ \boldsymbol{u} \in \boldsymbol{H}^1(\Omega) \,|\, \boldsymbol{u} = 0 \quad \text{on } \Gamma_\alpha \right\}.$$

In this work, the Navier-Stokes equations will be used in Chapters 1 and 2 to solve coupled fluid-structure systems. In order to derive the weak form of the Navier-Stokes system, we introduce the following continuous bilinear and trilinear forms

$$a(\boldsymbol{u},\boldsymbol{v}) = \frac{1}{2} \int_\Omega \left(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^T \right) : \left(\nabla\boldsymbol{v} + (\nabla\boldsymbol{v})^T \right) d\boldsymbol{x} \qquad \forall \boldsymbol{u},\boldsymbol{v} \in \boldsymbol{H}^1(\Omega),$$

$$b(\boldsymbol{u},\psi) = -\int_\Omega \psi \nabla\cdot\boldsymbol{u}\,d\boldsymbol{x} \qquad \forall\psi \in L_0^2(\Omega), \qquad \forall\boldsymbol{u} \in \boldsymbol{H}^1(\Omega),$$

$$c(\boldsymbol{w};\boldsymbol{u},\boldsymbol{v}) = \int_\Omega \boldsymbol{w}\cdot\nabla\boldsymbol{u}\cdot\boldsymbol{v}\,d\boldsymbol{x} \qquad \forall\boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in \boldsymbol{H}^1(\Omega).$$

These forms are continuous, namely there exist some constants $k_a, k_b, k_c \in \mathbb{R}^+$ such that

$$
\begin{aligned}
|a(\boldsymbol{u},\boldsymbol{v})| &\le k_a \|\boldsymbol{u}\|_1 \|\boldsymbol{v}\|_1 & \forall\boldsymbol{u},\boldsymbol{v} \in \boldsymbol{H}^1(\Omega), \\
|b(\boldsymbol{u},\psi)| &\le k_b \|\boldsymbol{u}\|_1 \|\psi\| & \forall\psi \in L_0^2(\Omega),\, \forall\boldsymbol{u} \in \boldsymbol{H}^1(\Omega), \\
|c(\boldsymbol{w};\boldsymbol{u},\boldsymbol{v})| &\le \|\boldsymbol{u}\|_1 \|\boldsymbol{v}\|_1 \|\boldsymbol{w}\|_1 & \forall\boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in \boldsymbol{H}^1(\Omega).
\end{aligned}
$$

For more details concerning the bilinear and trilinear forms, see [42, 43]. Moreover, for some constants $c_a, c_b \in \mathbb{R}^+$ the following coercivity properties hold

$$a(\boldsymbol{u}, \boldsymbol{u}) \geq c_a \|\boldsymbol{u}\|_1^2 \qquad \boldsymbol{u} \in \boldsymbol{H}_{\Gamma_\alpha}^1(\Omega) \,,$$

$$\inf_{p \in L_0^2(\Omega)} \sup_{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)} \frac{b(\boldsymbol{v}, p)}{\|\boldsymbol{v}\|_1 \, \|p\|} \geq c_b \,.$$

Interested reader on the coercivity properties can consult [44, 45, 44, 46].

We now consider a domain $\Omega^t \subset \mathbb{R}^3$ with boundary Γ^t, that consists of a fluid part Ω_f^t and a solid part Ω_s^t, so that $\Omega^t = \Omega_s^t \cup \Omega_f^t$ and $\Omega_s^t \cap \Omega_f^t = \Gamma_i^t$ at $t \in (0, T]$. The surface $\Gamma_i^t = \partial \Omega_s^t \cap \partial \Omega_f^t$ shared between the solid and the fluid is called *fluid-structure interface*. The outer boundary $\Gamma^t = \partial \Omega^t$ is then split into the solid boundary $\Gamma_s^t = \Gamma^t \cap \partial \Omega_s^t$ and the fluid one $\Gamma_f^t = \Gamma^t \cap \partial \Omega_f^t$. The solid and liquid reference undeformed configurations are denoted with Ω_s^0 and Ω_f^0, respectively.

1.3.2 The mathematical model for monolithic FSI

Considering the introduced notation for the spaces and the domains, we can now introduce the mathematical model for the coupled monolithic solver for fluid-structure interaction simulations.

The evolution of the solid and fluid domain $\hat{\Omega}_f$ and $\hat{\Omega}_s$ are defined by

$$\mathcal{X}^s : \hat{\Omega}_s \times \mathbb{R}^+ \to \mathbb{R}^3 \,,$$
$$\mathcal{A}^f : \hat{\Omega}_f \times \mathbb{R}^+ \to \mathbb{R}^3 \,,$$

such that the range of $\mathcal{X}^s(\cdot, t)$ and $\mathcal{A}^f(\cdot, t)$ define Ω_s^t and Ω_f^t, respectively. Note that $\mathcal{X}^s$ is the same mapping $\mathcal{X}$ introduced in the last section in order to describe the motion of the solid structure. In fact, $\mathcal{X}^s$ maps any material point $\hat{\boldsymbol{x}}_s$ from the given fixed reference configuration $\hat{\Omega}_s$ to the current solid material configuration Ω_s^t. The solid displacement is then defined as

$$\boldsymbol{\eta}(\hat{\boldsymbol{x}}_s, t) = \mathcal{X}(\hat{\boldsymbol{x}}_s, t) - \hat{\boldsymbol{x}}_s \,. \tag{1.56}$$

The mapping $\mathcal{A}^f$ is such that $\mathcal{A}^f(\hat{\boldsymbol{x}}_f, t) = \hat{\boldsymbol{x}}_f + \hat{\boldsymbol{\eta}}_f(\hat{\boldsymbol{x}}_f, t)$, where $\hat{\boldsymbol{\eta}}_f(\hat{\boldsymbol{x}}_f, t)$ is defined as an arbitrary extension operator over the fluid domain $\hat{\Omega}_f$ and given by

$$\hat{\boldsymbol{\eta}}_f(\hat{\boldsymbol{x}}_f, t) = \text{Ext}(\boldsymbol{\eta}|_{\hat{\Gamma}_0^i}) \quad \text{in} \quad \hat{\Omega}_f \,. \tag{1.57}$$

Following the work in [47], an harmonic or Laplace extension operator has been used to evaluate the fluid region displacement. The velocity $\hat{\boldsymbol{w}}_f$ is defined by

$$\hat{\boldsymbol{w}}_f = \frac{\partial \hat{\boldsymbol{\eta}}_f}{\partial t} \quad \text{in} \quad \hat{\Omega}_f . \tag{1.58}$$

This quantity represents the velocity in terms of the reference coordinate $\hat{\boldsymbol{x}}_f$. The behavior of the fluid is described by the Navier-Stokes equations for incompressible flows as

$$
\begin{aligned}
\rho_f \frac{\partial \boldsymbol{u}_f}{\partial t}\bigg|_{\tilde{\mathcal{A}}} + \rho_f \left(\boldsymbol{u}_f - \boldsymbol{w}_f \right) \cdot \nabla \boldsymbol{u}_f - \nabla \cdot \boldsymbol{\sigma}_f = \rho_f \boldsymbol{g} & \quad \text{in} \quad (0, T) \times \Omega_f^t , \\
\nabla \cdot \boldsymbol{u}_f = \boldsymbol{0} & \quad \text{in} \quad (0, T) \times \Omega_f^t , \\
\boldsymbol{u}_f|_{t=0} = \boldsymbol{u}_0 & \quad \text{in} \quad \hat{\Omega}_f , \\
\boldsymbol{u}_f|_{\Gamma_{t,D}^{1,f} \cup \Gamma_{t,D}^{2,f}} = \boldsymbol{g}_f & \quad \text{in} \quad (0, T), , \\
\boldsymbol{\sigma}_f \cdot \boldsymbol{n}_f|_{\Gamma_{t,N}^{1,f} \cup \Gamma_{t,}^{2,f}} = \boldsymbol{h}_f & \quad \text{in} \quad (0, T) ,
\end{aligned}
\tag{1.59}
$$

where $\boldsymbol{u}_f$ is the fluid velocity, ρ_f is the constant density, $\boldsymbol{g}$ is the gravity acceleration vector, $\tilde{\mathcal{A}}$ denotes the ALE application that maps the reference fluid configuration $\hat{\Omega}_f$ into the current fluid configuration Ω_f^t and $\boldsymbol{w}_f$ denotes the fluid domain velocity. Moreover $\boldsymbol{n}$ is the unit normal vector that points outward from the boundary $\partial \Omega_f^t$ and $\boldsymbol{g}_f$, $\boldsymbol{h}_f$, $\boldsymbol{u}_0$ are given data. The flow state variables in the incompressible case are the pressure p_f and the velocity $\boldsymbol{u}_f$. In the following, the contribution of external forces such as gravity is assumed to be negligible. The constitutive relation for the stress tensor in the Newtonian incompressible case reads

$$\boldsymbol{\sigma}_f = -p_f \boldsymbol{I} + \boldsymbol{\tau}_f = -p_f \boldsymbol{I} + 2\mu_f \boldsymbol{\epsilon} \left(\boldsymbol{u}_f \right) , \tag{1.60}$$

where μ_f is the dynamic viscosity of the fluid, p_f the Lagrange multiplier associated to the incompressibility constraint and $\boldsymbol{\epsilon} \left(\boldsymbol{u}_f \right)$ the strain rate tensor defined as

$$\boldsymbol{\epsilon} \left(\boldsymbol{u}_f \right) = \frac{1}{2} \left(\nabla \boldsymbol{u}_f + (\nabla \boldsymbol{u}_f)^T \right) . \tag{1.61}$$

On the other hand, the governing equations for structural mechanics are the following momentum equations

$$\rho_s \frac{\partial \boldsymbol{u}_s}{\partial t} - \nabla \cdot \boldsymbol{\sigma}_s(\boldsymbol{\eta}_s) = \boldsymbol{0} \quad \text{in} \quad \Omega_s^t , \tag{1.62}$$

where ρ_s is the density of the solid material, $\boldsymbol{u}_s$ is the velocity field and $\boldsymbol{\sigma}_s$ its Cauchy stress tensor, which is a function of the solid region displacement $\boldsymbol{\eta}_s$

as we have seen in the last sections. Since the constitutive law for the solid stress tensor is expressed in terms of displacements one must solve both the balance equations (1.62) and the simple kinematic relation

$$\boldsymbol{u}_s = \frac{\partial \boldsymbol{\eta}_s}{\partial t}\,.$$

For the reference configuration we can recall the right Cauchy-Green deformation tensor $\boldsymbol{C}$ as

$$C_{ij} = F_{ki}F_{kj} \qquad i,j = 1,2,3\,,$$

where $\boldsymbol{F}$ is the deformation gradient tensor defined in (1.4). It can be written as $\boldsymbol{F} = \boldsymbol{I} + \nabla \boldsymbol{\eta}_s$. In a similar way in the current configuration we can define the left Cauchy-Green deformation tensor $\boldsymbol{b}$ as

$$b_{ij} = F_{ik}F_{jk} \qquad i,j = 1,2,3\,.$$

According with this notation and recalling we can now express the Cauchy stress tensor $\boldsymbol{\sigma}_s$, as

$$\sigma_{s,ij} = \frac{2}{J_C}\left[b_{ij}\left(I\,b_{ij} - b_{im}b_{mj}\right)\frac{J_C\delta_{ij}}{2}\right]\begin{pmatrix}\frac{\partial W}{\partial I}\\[4pt]\frac{\partial W}{\partial II}\\[4pt]\frac{\partial W}{\partial J_C}\end{pmatrix}, \tag{1.63}$$

where $I = C_{ii}$, $II = 1/2\left(I - C_{ij}C_{ji}\right)$ are the first and second invariant of the right Cauchy-Green strain tensor $\boldsymbol{C}$ and J_C its determinant. The quantity $W = W(I, II, J_C)$ is the strain energy of the system which depends on the constitutive law of the considered material. For example for a Neo-Hookian material, with respect to the current configuration, by considering (1.51) it is possible to obtain the energy function as

$$W(I, J_C) = \frac{1}{2}\mu_s\left(J_C^{-2/3}\operatorname{tr}\boldsymbol{C} - 3\right) + \frac{1}{2}\left(\lambda + \frac{2}{3}\mu_s\right)\left(\frac{1}{2}(J_C^2 - 1) - \ln J_C\right).$$

In the case of incompressible solid, $J_C = 1$ holds. Therefore the energy density function becomes

$$W(I, J_C) = \frac{1}{2}\mu_s\left(\operatorname{tr}\boldsymbol{C} - 3\right) \tag{1.64}$$

and the Cauchy stress tensor is defined by

$$\boldsymbol{\sigma}_s = -p_s\boldsymbol{I} + \boldsymbol{\sigma}_{s*}\,, \tag{1.65}$$

where $\boldsymbol{\sigma}_{s*}$ is the tensor obtained by using the equations (1.63) and (1.64). We shall now impose a set of boundary conditions at the interface Γ_t^i in order to solve the problem defined by (1.59)-(1.62). The coupling between the fluid and the solid model determines such a boundary condition. It consists of imposing the continuity of velocity and stress at the interface Γ_i^t as

$$\boldsymbol{u}_f|_{\Gamma_t^i} = \boldsymbol{u}_s|_{\Gamma_t^i}\,, \tag{1.66}$$

$$\boldsymbol{\sigma}_f \cdot \boldsymbol{n}_f|_{\Gamma_t^i} + \boldsymbol{\sigma}_s \cdot \boldsymbol{n}_s|_{\Gamma_t^i} = \boldsymbol{0}\,. \tag{1.67}$$

We introduce now some functional space useful in the rest of this section to describe the weak formulation of the studied problem.

$$\begin{aligned}
\boldsymbol{V}_t &= \{\boldsymbol{\phi} \in \boldsymbol{H}^1(\Omega_f^t) : \boldsymbol{\phi}|_{\Gamma_{t,D}^{1,f} \cup \Gamma_{t,D}^{2,f}} = \boldsymbol{0}\}\,, \\
\boldsymbol{V}_g^t &= \{\boldsymbol{\phi} \in \boldsymbol{H}^1(\Omega_f^t) : \boldsymbol{\phi}|_{\Gamma_{t,D}^{1,f} \cup \Gamma_{t,D}^{2,f}} = \boldsymbol{g}^f\}\,, \\
Q^t &= L^2(\Omega_f^t)\,, \\
\boldsymbol{M}^0 &= \{\boldsymbol{\psi} \in \boldsymbol{H}^1(\hat{\Omega}_s) : \boldsymbol{\psi}|_{\hat{\Gamma}_{0,D}^{1,s} \cup \hat{\Gamma}_{0,D}^{2,s} \cup \hat{\Gamma}_{0,D}^3} = \boldsymbol{0}\}\,, \\
\boldsymbol{M}_g^0 &= \{\boldsymbol{\psi} \in \boldsymbol{H}^1(\hat{\Omega}_s) : \boldsymbol{\psi}|_{\hat{\Gamma}_{0,D}^{1,s} \cup \hat{\Gamma}_{0,D}^{2,s} \cup \hat{\Gamma}_{0,D}^3} = \boldsymbol{g}_s\}\,, \\
D^0 &= L^2(\hat{\Omega}_s)\,.
\end{aligned}$$

In addition, let us introduce the following bilinear form

$$a_f(\boldsymbol{u}_f, \boldsymbol{\phi}) = \int_{\Omega_f} \boldsymbol{\tau}_f(\boldsymbol{u}_f) : \nabla\boldsymbol{\phi}\,d\mathbf{x} = \mu(\nabla\boldsymbol{u}_f + (\nabla\boldsymbol{u}_f)^T, \nabla\boldsymbol{\phi})\,, \tag{1.68}$$

where we denote with $\boldsymbol{\tau}_f$ the fluid viscosity tensor. The variational formulation of the presented problem follows standard techniques. We multiply the first equation of (1.59) by appropriate test functions, performing integrations on the whole domain and keeping into account the boundary and interface conditions. By doing so, we find the following fluid momentum equation for the unknowns $\boldsymbol{u} \in \boldsymbol{V}_g^t$ and $p \in Q^t$

$$\begin{aligned}
\rho_f &\left(\left.\frac{\partial \boldsymbol{u}_f}{\partial t}\right|_{\tilde{\mathcal{A}}}, \boldsymbol{\phi}\right) + \rho_f((\boldsymbol{u}_f - \boldsymbol{w}_f) \cdot \nabla\boldsymbol{u}_f, \boldsymbol{\phi}) + a(\boldsymbol{u}_f, \boldsymbol{\phi}) \\
&= (p_f, \nabla \cdot \boldsymbol{\phi}) + \int_{\Gamma_i^t} (\boldsymbol{\sigma}_f \cdot \boldsymbol{n}_f) \cdot \boldsymbol{\phi}\,d\boldsymbol{\gamma} + \int_{\Gamma_N^f} \boldsymbol{h}^f \cdot \boldsymbol{\phi}\,d\boldsymbol{\gamma}\,, \\
(q, &\nabla \cdot \boldsymbol{u}_f) = 0\,, \\
\boldsymbol{u}_f&|_{t=0} = \boldsymbol{u}_0^f\,,
\end{aligned} \tag{1.69}$$

for all $\boldsymbol{\phi} \in \boldsymbol{V}^t$ and $q \in Q^t$. In a similar way, we define the following bilinear form

$$a_s(\boldsymbol{\eta}_s, \boldsymbol{\psi}) = (\boldsymbol{\sigma}_s(\boldsymbol{u_s}), \nabla \boldsymbol{\psi}) \,. \tag{1.70}$$

Considering a displacement $\boldsymbol{\eta}_s \circ \mathcal{X}_s \in \boldsymbol{M}_g^0$ and pressure $p_s \circ \mathcal{X}_s \in \boldsymbol{D}_0$, we can now obtain the weak formulation for the solid sub-problem

$$
\rho_s \left(\frac{\partial^2}{\partial t^2} \boldsymbol{\eta}_s, \boldsymbol{\psi} \right) + a_s(\boldsymbol{\eta}_s, \boldsymbol{\psi}) - (p_s, \nabla \cdot \boldsymbol{\psi}) = \int_{\Gamma_t^i} (\boldsymbol{\sigma}_s \cdot \boldsymbol{n}_s) \cdot \boldsymbol{\psi} \, d\boldsymbol{\gamma}
$$

$$
+ \int_{\Gamma_N^s} \boldsymbol{h}_s \cdot \boldsymbol{\psi} \, d\boldsymbol{\gamma} \,, \tag{1.71}
$$

$$(d, \nabla \cdot \boldsymbol{\eta}_s) = 0 \,,$$

$$\boldsymbol{\eta}_s|_{t=0} = \boldsymbol{\eta}_{s,0} \,, \qquad \boldsymbol{u}_s|_{t=0} = \boldsymbol{u}_{s,0} \,,$$

for all $\boldsymbol{\psi} \circ \mathcal{X}_s \in \boldsymbol{M}^0$ and $d \circ \mathcal{X}_s \in \boldsymbol{D}^0$. Let us introduce a global weak formulation for the fluid-structure problem. If we define the functional space

$$\boldsymbol{S}^t = \{ (\boldsymbol{\phi}, \boldsymbol{\psi} \circ \mathcal{X}_s) \in \boldsymbol{V}^t \times \boldsymbol{M}^0 : \boldsymbol{\psi}|_{\Gamma_t^i} = \boldsymbol{\phi}|_{\Gamma_t^i} \} \,, \tag{1.72}$$

from (1.66), (1.67), (1.69) and (1.71), we can write the FSI problem in the coupled formulation as

$$
\rho_f \left(\frac{\partial \boldsymbol{u}_f}{\partial t} \Big|_{\tilde{\mathcal{A}}}, \boldsymbol{\varphi} \right) + \rho_f((\boldsymbol{u}_f - \boldsymbol{w}_f) \cdot \nabla \boldsymbol{u}_f, \boldsymbol{\varphi}) + a(\boldsymbol{u}_f, \boldsymbol{\varphi})
$$

$$
- (p_f, \nabla \cdot \boldsymbol{\varphi}) + \rho_s \left(\frac{\partial^2}{\partial t^2} \boldsymbol{\eta}_s, \boldsymbol{\varphi} \right) + a_s(\boldsymbol{\eta}_s, \boldsymbol{\varphi})
$$

$$
- (p_s, \nabla \cdot \boldsymbol{\varphi}) - \int_{\Gamma_N^s} \boldsymbol{h}_s \cdot \boldsymbol{\varphi} \, d\boldsymbol{\gamma} - \int_{\Gamma_N^f} \boldsymbol{h}_f \cdot \boldsymbol{\varphi} \, d\boldsymbol{\gamma} = 0 \,, \tag{1.73}
$$

$$(q, \nabla \cdot \boldsymbol{u}_f) = 0 \qquad (d, \nabla \cdot \boldsymbol{\eta}_s) = 0 \,,$$

$$\boldsymbol{u}_f|_{t=0} = \boldsymbol{u}_0^f \qquad \boldsymbol{\eta}_s|_{t=0} = \boldsymbol{\eta}_{s,0} \qquad \boldsymbol{u}_s|_{t=0} = \boldsymbol{u}_0^s \,,$$

for all $\boldsymbol{\varphi} \in \boldsymbol{S}^t$. It is worth noting that by using the coupling conditions (1.66), (1.67) and this particular choice of the fluid-structure test functions, the boundary terms that appear in the fluid-solid interface Γ_t^i cancel out. This assures that forces at the interface are always computed in an exact way. The numerical simulations of the fluid-structure interaction problems can be carried out through the discretization of the system (1.73) in the finite element code FEMuS.

1.4   A Koiter shell model applied to FSI

In literature many works can be found on the reduction of the computational cost of fluid-structure interaction simulations. These works are generally based on velocity-pressure splitting preconditioners, that keep the original number of degrees of freedom and preserve exact boundary conditions, see e.g. [48, 49, 50]. In the framework of the reduction of computational costs of FSI simulations, in this section a technique for the reduction of the dimensionality of the solid structure is shown. A model based on the Koiter shell equations is used to this scope [4]. In order to couple the fluid and the structure domains, the Koiter shell equations are embedded into the fluid equations as a Robin boundary condition [5]. In order to preserve the stability of the numerical scheme, the coupling fluid-structure conditions are automatically treated in an implicit way. This model has many applications in cases where a fluid interacts with a thin membrane that deforms mainly in the normal direction. It is particularly used for hemodynamic applications.

This section is organized as follows. The Koiter shell model is firstly presented. Then, the presented structure model is embedded in the fluid equation in order to have the full fluid-structure system.

1.4.1   The linear Koiter shell model

The introduction of the linear Koiter model will follow the mathematical formulation presented in [51]. The Koiter shell approach relies on the assumptions that the structure displacements are small and normal to the surface of the shell. As reported in Figure 1.1, let $\boldsymbol{\theta}(\boldsymbol{x})$ be a mapping defined as

$$\boldsymbol{\theta}(\boldsymbol{x}) \,:\, \omega \subset \mathbb{R}^2 \to \Gamma_t \subset \mathbb{R}^3 \,,$$

where ω is the undeformed reference membrane and Γ_t is the deformed membrane. We now consider the tangential base $\boldsymbol{a}_\alpha = \dfrac{\partial \boldsymbol{r}}{\partial \boldsymbol{x}_\alpha}$, for $\alpha = 1, 2$. Then, $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ define the tangential plane to the reference shell, while $\boldsymbol{a}_n = (\boldsymbol{a}_1 \times \boldsymbol{a}_2)/|\boldsymbol{a}_1 \times \boldsymbol{a}_2|$ defines the unit vector normal to the shell surface. We define now the covariant components of the metric tensor of the middle deformed surface as $A_{\alpha\beta} = \boldsymbol{a}_\alpha \cdot \boldsymbol{a}_\beta$. Another fundamental tensor is

$$B_{\alpha\beta} = \boldsymbol{a}_n \cdot \frac{\partial}{\partial x_\alpha}\left(\frac{\partial \boldsymbol{r}}{\partial x_\beta}\right) = \boldsymbol{a}_n \cdot \frac{\partial \boldsymbol{a}_\alpha}{\partial x_\beta} \,.$$
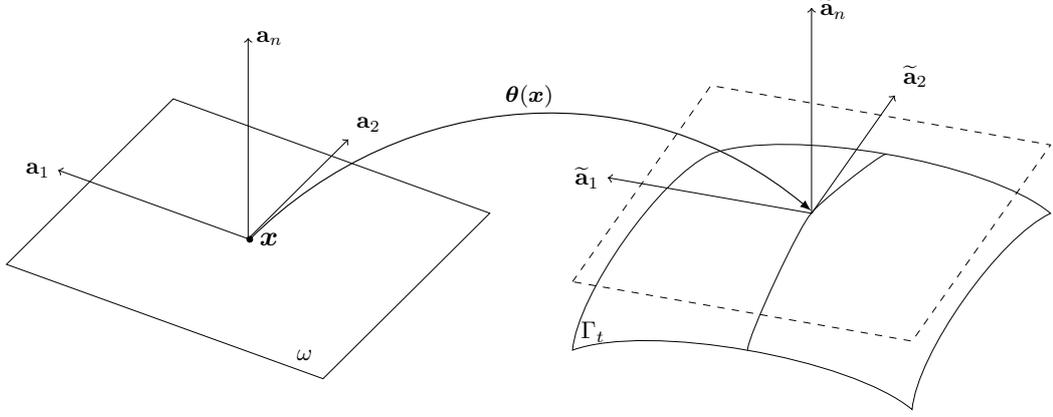
Figure 1.1: Regular mapping to identify the reference shell surface ω.

In the following, we will use the notation $A^{\alpha\beta} = (A_{\alpha\beta})^{-1}$. We now introduce the strain measures.

Theorem 1.3. *Let consider the domain $\omega \subset \mathbb{R}^2$. Let also consider the immersion $\boldsymbol{\theta} \in C^2(\bar{\omega}, \mathbb{R}^3)$. For a given displacement $\boldsymbol{\eta} = \eta_i \boldsymbol{a}^i$ of the surface ω, defined by the immersion $\omega = \boldsymbol{\theta}(\bar{\omega})$, with smooth enough covariant components $\eta_i : \bar{\omega} \to \mathbb{R}$, let the change of metric tensor $\gamma_{\alpha\beta}(\boldsymbol{\eta}) : \bar{\omega} \to \mathbb{R}$ be defined as*

$$\gamma_{\alpha\beta}(\boldsymbol{\eta}) = \frac{1}{2}(\tilde{A}_{\alpha\beta}(\boldsymbol{\eta}) - A_{\alpha\beta}), \tag{1.74}$$

where $\tilde{A}_{\alpha\beta}(\boldsymbol{\eta})$ is the covariant component of the first fundamental form of the surface $(\boldsymbol{\theta} + \boldsymbol{\eta})(\bar{\omega})$. $A_{\alpha\beta}$, as introduced above, is the covariant component of the first fundamental form of the surface $\boldsymbol{\theta}(\bar{\omega})$. Then, by considering only the linear part with respect to $\boldsymbol{\eta}$ of (1.74), we obtain

$$
\begin{aligned}
\gamma_{\alpha\beta}(\boldsymbol{\eta}) &= \frac{1}{2}(\partial_\beta \boldsymbol{\eta} \cdot \boldsymbol{a}_\alpha + \partial_\alpha \boldsymbol{\eta} \cdot \boldsymbol{a}_\beta) \\
&= \frac{1}{2}(\eta_{\alpha|\beta} + \eta_{\beta|\alpha}) - B_{\alpha\beta}\eta_n \\
&= \frac{1}{2}(\partial_\beta \eta_\alpha + \partial_\alpha \eta_\beta) - \Gamma_{\alpha\beta}^k \eta_k - B_{\alpha\beta}\eta_n\,,
\end{aligned}
\tag{1.75}
$$

where the covariant derivatives are defined by $\eta_{\alpha|\beta} = \partial_\beta \eta_\alpha - \Gamma_{\alpha\beta}^k \eta_k$, and $\Gamma_{\alpha\beta}^k$ are the Christoffel symbols, defined by $\Gamma_{\alpha\beta}^k = \boldsymbol{a}^k \partial_\alpha \boldsymbol{a}_\beta$ (see [52]). Note that $\gamma_{\alpha\beta}(\boldsymbol{\eta}) = \gamma_{\beta\alpha}(\boldsymbol{\eta})$. Thus,

$$\eta_\alpha \in H^1(\omega)\,,\ \eta_n \in L^2(\omega) \Rightarrow \gamma_{\alpha\beta}(\boldsymbol{\eta}) \in L^2(\omega)$$

Proof. The covariant components can be written as

$$\tilde{A}_{\alpha\beta}(\boldsymbol{\eta}) = \partial_\alpha(\boldsymbol{\theta} + \boldsymbol{\eta}) \cdot \partial_\beta(\boldsymbol{\theta} + \boldsymbol{\eta}) \,. \tag{1.76}$$

The we can use the relation $\partial_\alpha(\boldsymbol{\theta} + \boldsymbol{\eta}) = \boldsymbol{a}_\alpha + \partial_\alpha\boldsymbol{\eta}$ to show that (1.76) can be written as

$$\tilde{A}_{\alpha\beta}(\boldsymbol{\eta}) = (\boldsymbol{a}_\alpha + \partial_\alpha\boldsymbol{\eta}) \cdot (\boldsymbol{a}_\beta + \partial_\beta\boldsymbol{\eta}) = A_{\alpha\beta}(\boldsymbol{\eta}) + \partial_\beta\boldsymbol{\eta} \cdot \boldsymbol{a}_\beta + \partial_\alpha\boldsymbol{\eta} \cdot \partial_\beta\boldsymbol{\eta} \,,$$

therefore

$$\gamma_{\alpha\beta}(\boldsymbol{\eta}) = \frac{1}{2}[\tilde{A}_{\alpha\beta}(\boldsymbol{\eta}) - A_{\alpha\beta}]^{\mathrm{lin}} = \frac{1}{2}(\partial_\beta\boldsymbol{\eta} \cdot \boldsymbol{a}_\alpha + \partial_\alpha\boldsymbol{\eta} \cdot \boldsymbol{a}_\beta) \,.$$

In order to prove the other relations of equation (1.75), we consider now the Gauss and Weingarten equations

$$\partial_\alpha\boldsymbol{a}_\beta = \Gamma^k_{\alpha\beta}\boldsymbol{a}_k + B_{\alpha\beta}\boldsymbol{a}^n \,,$$

$$\partial_\alpha\boldsymbol{a}_n = -B^k_\alpha\boldsymbol{a}_k \,,$$

that imply

$$\partial_\alpha\boldsymbol{\eta} = \partial_\alpha(\eta_i\boldsymbol{a}^i) = (\partial_\alpha\eta_k - \Gamma^\tau_{\alpha k}\eta_\tau - B_{\alpha k}\eta_n)\boldsymbol{a}^k + (\partial_\alpha\eta_n + B^\tau_\alpha\eta_\tau)\boldsymbol{a}^n \,.$$

Therefore, we obtain

$$\partial_\alpha\boldsymbol{a}^\tau = -\Gamma^\tau_{\alpha k}\boldsymbol{a}^k + B^\tau_\alpha\boldsymbol{a}^n$$

$$\partial_\alpha\boldsymbol{a}^n = -B_{\alpha k}\boldsymbol{a}^k \,.$$

$\square$

Note that the introduced tensor $\gamma_{\alpha\beta}(\boldsymbol{\eta})$ is associated with the displacement of the surface S, as well as the covariant components of the linearized change of curvature tensor $\rho_{\alpha\beta}$ that will be introduced in the next theorem.

Theorem 1.4. *Consider the hypotheses of the Theorem 1.3, with the only difference $\boldsymbol{\theta} \in C^3(\bar{\omega}; \mathbb{R}^3)$. Let the functions $\rho_{\alpha\beta}(\boldsymbol{\eta}) : \bar{\omega} \to \mathbb{R}$ be defined by*

$$\rho_{\alpha\beta} = [\tilde{B}_{\alpha\beta}(\boldsymbol{\eta}) - B_{\alpha\beta}] \,,$$

where $\tilde{B}_{\alpha\beta}(\boldsymbol{\eta})$ is the covariant component of the second fundamental form of the surface $(\boldsymbol{\theta} + \boldsymbol{\eta})(\bar{\omega})$. $B_{\alpha\beta}$, as introduced above, is the covariant component of the second fundamental form of the surface $\boldsymbol{\theta}(\bar{\omega})$. Again, we consider only

the linear part with respect to $\boldsymbol{\eta}$ of the introduced expression. Therefore, if we define the covariant derivatives as

$$\eta_{\alpha|\beta} = \partial_\beta \eta_\alpha - \Gamma^k_{\alpha\beta} \eta_k\,,$$
$$\eta_{n|\alpha\beta} = \partial_{\alpha\beta} \eta_n - \Gamma^k_{\alpha\beta} \partial_k \eta_n\,,$$
$$B^\tau_\beta|_\alpha = \partial_\alpha B^\tau_\beta + \Gamma^\tau_{\alpha k} B^k_\beta - \Gamma^k_{\alpha\beta} B^\tau_k\,,$$

we obtain

$$
\begin{aligned}
\rho_{\alpha\beta}(\boldsymbol{\eta}) &= (\partial_{\alpha\beta}\boldsymbol{\eta} - \Gamma^k_{\alpha\beta}\partial_k\boldsymbol{\eta}) \cdot \boldsymbol{a}_n \\
&= \eta_{n|\alpha\beta} - B^k_\alpha B_{k\beta}\eta_n + B^k_\alpha \eta_{k|\beta} + B^\tau_\beta \eta_{\tau|\alpha} + B^\tau_\beta|_\alpha \eta_\tau \\
&= \partial_{\alpha\beta}\eta_n - \Gamma^k_{\alpha\beta}\partial_k\eta_n - B^k_\alpha B_{k\beta}\eta_n + B^k_\alpha(\partial_\beta\eta_k - \Gamma^\tau_{\beta k}\eta_\tau) \\
&\quad + B^\tau_\beta(\partial_\alpha\eta_\tau - \Gamma^k_{\alpha\tau}\eta_k) + (\partial_\alpha B^\tau_\beta + \Gamma^\tau_{\alpha k}B^k_\beta - \Gamma^k_{\alpha\beta}B^\tau_k)\eta_\tau\,.
\end{aligned}
$$

Then,

$$\eta_\alpha \in H^1(\omega)\,,\ \eta_n \in H^2(\omega) \Rightarrow \rho_{\alpha\beta} \in L^2(\omega)\,.$$

In addition, both $\rho_{\alpha\beta}$ and $B^\tau_\beta|_\alpha$ satisfy the symmetry relations $\rho_{\alpha\beta} = \rho_{\beta\alpha}$ and $B^\tau_\beta|_\alpha = B^\tau_\alpha|_\beta$

The proof of the Theorem 1.4 is not reported in this work. Interested reader can find it in [51].

We can now introduce the linear Koiter shell equations. Let ω be the considered shell domain, and let υ_s be a measurable subset of $\partial\omega$. In the following, ∂_t will denote the outer normal derivative operator along ω. Since $\gamma_{\alpha\beta}(\boldsymbol{\eta})$, $\rho_{\alpha\beta}(\boldsymbol{\eta}) \in L^2(\omega)$, we can introduce the functional space $\boldsymbol{V}(\omega)$ such that

$$\boldsymbol{V}(\omega) = \{\boldsymbol{\eta} = \eta_i \boldsymbol{a}^i\ :\ \eta_\alpha \in H^1(\omega)\,,\ \eta_n \in H^2(\omega)\,,\ \eta_i = \partial_t\eta_n = 0 \text{ on } \upsilon_s\}\,.$$

For more information about the introduced functional spaces and the derivation of the used equations see [51]. Tthe unknown displacement field $\boldsymbol{\xi}^\varepsilon = \xi^\varepsilon_1 \boldsymbol{a}^i$ of the middle surface of the shell should be a stationary point over the space $\boldsymbol{V}(\omega)$ of the energy functional defined as

$$
\begin{aligned}
\mathcal{K}(\boldsymbol{\eta}) = \frac{1}{2}\int_\omega \Big(&\varepsilon A^{\alpha\beta\sigma\tau}\gamma_{\sigma\tau}(\boldsymbol{\eta})\gamma_{\alpha\beta}(\boldsymbol{\eta}) + \frac{\varepsilon^3}{3}A^{\alpha\beta\sigma\tau}\rho_{\sigma\tau}(\boldsymbol{\eta})\rho_{\alpha\beta}(\boldsymbol{\eta})\Big)\sqrt{a}\,dy \\
&- \int_\omega f^{i,\varepsilon}\eta_i\sqrt{a}\,dy \qquad\qquad \forall\boldsymbol{\eta} \in \boldsymbol{V}(\omega)\,,
\end{aligned}
$$

(1.77)

where $a = a(\boldsymbol{\eta}) = \det(A_{\alpha\beta}(\boldsymbol{\eta}))$. The contravariant components of the shell elasticity tensor $A^{\alpha\beta\sigma\tau}$ is defined as

$$A^{\alpha\beta\sigma\tau} = \frac{2E\nu}{1-\nu^2} A^{\alpha\beta} A^{\sigma\tau} + \frac{E}{1+\nu}(A^{\alpha\sigma} A^{\beta\tau} + A^{\alpha\tau} A^{\beta\sigma}) \,,$$

where ν is the Poisson coefficient and E Young modulus of the solid material. The given functions $f^{i,\varepsilon} \in L^2(\omega)$ take into account the forces applied to the shell. Note also that the boundary conditions $\eta_i = \partial_t \eta_n = 0$ on v_s means that the shell is clamped along the boundaries of its middle surface. We can now derive the variational equation for the vector field $\boldsymbol{\eta}$ as

$$\int_\omega \left(\varepsilon A^{\alpha\beta\sigma\tau} \gamma_{\sigma\tau}(\boldsymbol{\xi}^\varepsilon) \gamma_{\alpha\beta}(\boldsymbol{\eta}) + \frac{\varepsilon^3}{3} A^{\alpha\beta\sigma\tau} \rho_{\sigma\tau}(\boldsymbol{\xi}^\varepsilon) \rho_{\alpha\beta}(\boldsymbol{\eta}) \right) \sqrt{a}\, dy$$
$$= \int_\omega f^{i,\varepsilon} \eta_i \sqrt{a}\, dy \,, \qquad \forall \boldsymbol{\xi}^\varepsilon \in \boldsymbol{V}(\omega) \,. \tag{1.78}$$

The model used in this work is based on some simplifying assumptions: small deformations of the solid shell, negligible bending terms, only normal displacement, and isotropic and homogeneous material. In the following, the reference domain of the shell structure ω will be denoted with Γ_s. Under these assumptions, it is possible to obtain the following model from (1.78) (see [5])

$$\int_{\Gamma_s} \rho_s \varepsilon \frac{\partial^2 \boldsymbol{\eta}}{\partial t^2}\, d\boldsymbol{x} + \int_{\Gamma_s} \varepsilon A^{\alpha\beta\sigma\tau} \gamma_{\alpha\beta}(\boldsymbol{\eta}) \gamma_{\lambda\delta}(\boldsymbol{\psi})\, d\Gamma = \int_{\Gamma_s} \boldsymbol{f}_s \cdot \boldsymbol{\psi}\, d\Gamma \,, \tag{1.79}$$

where the displacement $\boldsymbol{\xi}^\varepsilon$ has been replaced with the test function notation $\boldsymbol{\psi}$. The density of the solid shell is denoted by ρ_s.

If we restrict the membrane displacements only to normal direction, then we can further simplify the model (1.79), and reduce it to a simple scalar equation for η_n. In strong form we have

$$\rho_s \varepsilon \frac{\partial^2 \eta_n}{\partial t^2} + \beta \eta_n = f_s \qquad \text{on} \quad \Gamma_s \,,$$
$$\text{with} \quad \eta_n|_{t=0} = \eta_0 \,, \quad \left. \frac{\partial \eta_n}{\partial t} \right|_{t=0} = \eta_v \quad \text{on} \quad \Gamma_s \,, \tag{1.80}$$

where

$$\beta(x_1, x_2) = \frac{\varepsilon E}{1-\nu^2} \left((1-\nu) B_\beta^k B_k^\beta + \nu B_\beta^\beta B_k^k \right) . \tag{1.81}$$

The prestressed term

The presented Koiter model does not account for prestressed loading along the shell structure. Let consider the deformed non-shell configuration Ω_s. Note that Ω_s has thickness ε, and the shell surface is defined as the middle surface of it. In weak formulation, the prestress term reads

$$\int_{\Omega_s} \nabla \boldsymbol{\eta} P : \nabla \boldsymbol{\xi}^\varepsilon \, d\boldsymbol{x} \,, \tag{1.82}$$

where P is the Cauchy stress tensor in the deformed configuration for only tangential stresses in Ω_s [53]. In the rest of this section, we will follow the procedure presented in [5]. Now we want to lead back the equation (1.82) to the membrane case, by taking the limit for $\varepsilon \to 0$. We can write the deformation field as

$$\boldsymbol{\eta} = \eta_i(x_1, x_2) a^i + x_n \Big(\frac{\partial \eta_n}{\partial x_\alpha} + B_\alpha^k \eta_k \Big) a^\alpha \,.$$

The terms where x_n appears are of higher order in ε, so we can neglect them. We introduce now the surface covariant derivative of a vector field (see [52]). The covariant derivative of $\boldsymbol{\eta}$ is defined as

$$\boldsymbol{\eta}_{\alpha|k}^s = \frac{\partial \eta_\alpha}{\partial x_k} - \Gamma_{\alpha k}^\beta \eta_\beta \,, \qquad \text{with } \Gamma_{\alpha k}^\beta = a^\beta \cdot \frac{\partial a_k}{\partial x_\alpha} \,.$$

With this notation, we can write the three-dimensional covariant derivatives of $\boldsymbol{\eta}$ as

$$\boldsymbol{\eta}_{\alpha|\beta} = \boldsymbol{\eta}_{\alpha|\beta}^s - B_{\alpha\beta}\eta_n \,, \qquad\qquad \boldsymbol{\eta}_{\alpha|n} = -\Big(\frac{\partial \eta_n}{\partial x_\alpha} + B_\alpha^k \eta_k \Big) \,,$$

$$\boldsymbol{\eta}_{n|\alpha} = \frac{\partial \eta_n}{\partial x_\alpha} + B_\alpha^k \eta_k \,, \qquad\qquad \boldsymbol{\eta}_{n|n} = 0 \,.$$

Now by integrating (1.82) considering $\Omega_s = \Gamma_s \times [-\varepsilon/2, +\varepsilon/2]$ we have

$$\begin{aligned}
\int_{\Omega_s} \boldsymbol{\eta}_{k|\alpha} P^{\alpha\beta} \boldsymbol{\xi}_{k|\beta}^\varepsilon \, d\boldsymbol{x} &= \int_{\Gamma_s} \int_{-\varepsilon/2}^{\varepsilon/2} B_{k\alpha} \eta_n P^{\alpha\beta} B_{k\beta} \xi_n^\varepsilon \, dx \, d\Gamma \\
&= \int_{\Gamma_s} \varepsilon B_{k\alpha} P^{\alpha\beta} B_{k\beta} \eta_n \xi_n^\varepsilon \, d\Gamma \,,
\end{aligned} \tag{1.83}$$

and

$$\begin{aligned}
\int_{\Omega_s} \boldsymbol{\eta}_{n|\alpha} P^{\alpha\beta} \boldsymbol{\xi}_{n|\beta}^\varepsilon \, d\boldsymbol{x} &= \int_{\Gamma_s} \int_{-\varepsilon/2}^{\varepsilon/2} \frac{\partial \eta_n}{\partial x_\alpha} P^{\alpha\beta} \frac{\partial \xi_n^\varepsilon}{\partial x_\beta} \, dx \, d\Gamma \\
&= \int_{\Gamma_s} \varepsilon P^{\alpha\beta} \frac{\partial \eta_n}{\partial x_\alpha} \frac{\partial \xi_n^\varepsilon}{\partial x_\beta} \, d\Gamma \,.
\end{aligned} \tag{1.84}$$

The equation (1.83) can be incorporated into the coefficient β introduced in (1.81). The equation (1.84) gives a second derivative in space to be added inside the model (1.80), obtaining

$$
\begin{aligned}
&\rho_s \varepsilon \frac{\partial^2 \eta_n}{\partial t^2} + \beta^* \eta_n - \nabla \cdot (P \nabla \eta_n) = f_s \quad \text{on} \quad \Gamma_s \ , \\
&\text{with} \quad \eta_n|_{t=0} = \eta_0 \ , \quad \left.\frac{\partial \eta_n}{\partial t}\right|_{t=0} = \eta_v \quad \text{on} \quad \Gamma_s \ ,
\end{aligned}
\tag{1.85}
$$

and

$$
\beta^*(x_1, x_2) = \varepsilon \left(\frac{E}{1 - \nu^2} \left((1 - \nu) B_\beta^k B_k^\beta + \nu B_\beta^\beta B_k^k \right) + B_{k\alpha} P^{\alpha\beta} B_{k\beta} \right) .
$$

The model (1.85) must be completed with proper boundary conditions, e.g. $\eta_n|_{\partial \Gamma_s} = 0$. Note that the presented prestressed model can be used when the deformed configuration is close enough to the reference one, in order to consider an isotropic elastic tensor.

The cylindrical geometry

Let now consider a cylindrical geometry, in order to show how to explicitly calculate all the introduced tensors in a simple geometry. Moreover, many of the results presented at the end of this chapter are based on a cylindrical domain. If we consider a system of cylindrical coordinates and a cylinder of radius R, we have $r(\theta, z) = \{(x, y, z) \in \mathbb{R}^3 \mid x = R\cos(\theta), y = R\sin(\theta), z = z\}$. The covariant basis is given by

$$
\boldsymbol{a}_1 = \begin{bmatrix} -R\sin\theta \\ R\cos\theta \\ 0 \end{bmatrix} , \quad
\boldsymbol{a}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} , \quad
\boldsymbol{a}_n = \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \end{bmatrix} ,
$$

this implies that

$$
A_{\alpha\beta} = \begin{bmatrix} R^2 & 0 \\ 0 & 1 \end{bmatrix} , \quad
B_{\alpha\beta} = \begin{bmatrix} -R & 0 \\ 0 & 0 \end{bmatrix} ,
$$

therefore

$$
B_\beta^\alpha = A^{\alpha k} B_{k\beta} = \begin{bmatrix} 1/R^2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -R & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1/R & 0 \\ 0 & 0 \end{bmatrix} .
$$

Then, it is easy to show that

$$\beta = \frac{\varepsilon E}{1 - \nu^2}\big((1 - \nu)B_\beta^k B_k^\beta + \nu B_\beta^\beta B_k^k\big) = \frac{\varepsilon E}{1 - \nu^2}\Big(\frac{1 - \nu}{R^2} + \frac{\nu}{R^2}\Big)$$
$$= \frac{\varepsilon E}{1 - \nu^2}\frac{1}{R^2}\,. \tag{1.86}$$

Moreover, since the prestress term acts only in the longitudinal dimension, in the cylindrical case we have

$$P = \begin{bmatrix} 0 & 0 \\ 0 & P^{zz} \end{bmatrix}\,.$$

In this case the system (1.85) can be simplified as a one dimensional equation

$$\rho_s \varepsilon \frac{\partial^2 \eta_n}{\partial t^2} + \beta \eta_n - \mu_s \frac{\partial^2 \eta_n}{\partial z^2} = f_s\,, \tag{1.87}$$

where $\mu_s = P^{zz}$. Interested reader can consult [54, 55] for more information on the applications of this model to hemodynamic.

1.4.2 The coupled fluid-shell problem

In this section the shell model presented in the previous section is embedded inside the fluid equations to properly simulate the fluid-structure interaction system. Now, we first introduce the model for the fluid equations. Let consider a Newtonian incompressible fluid flowing through a deformable domain $\Omega_f = \Omega_f(t)$. The domain is not fixed in time, because of the nature of the solved physical problem. In the following, the time dependence of Ω_f will be omitted for notation simplicity.

As we have seen in the previous section, the fluid is modeled following the ALE formulation [2, 33, 56] as

$$\rho_f \frac{\partial \boldsymbol{u}}{\partial t}\Big|_{\mathcal{A}} + \rho_f[(\boldsymbol{u} - \boldsymbol{w}) \cdot \nabla]\boldsymbol{u} - \nabla \cdot \boldsymbol{\sigma}^f = \boldsymbol{0} \qquad \text{on } \Omega_f\,, \tag{1.88}$$

$$\nabla \cdot \boldsymbol{u} = 0 \qquad \text{on } \Omega_f\,, \tag{1.89}$$

where ρ_f and $\boldsymbol{u}$ are the density and the velocity vector of the fluid, and $\boldsymbol{\sigma}^f$ is the Cauchy stress tensor of the fluid written as $\boldsymbol{\sigma}^f = -p\boldsymbol{I} + \mu(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^T)$, where p and μ are the pressure and the dynamic viscosity of the fluid, respectively. The system of equations (1.88) is completed with appropriate

boundary conditions. In this framework is important to compute properly the value of the velocity of the points of the fluid domain $\boldsymbol{w}$. On the moving boundaries we have

$$\boldsymbol{w}|_{\Gamma_s} = \boldsymbol{u}|_\Gamma \circ r(\boldsymbol{x}) \,.$$

Moreover, on the other boundaries we have in the direction normal to the surface

$$(\boldsymbol{w} \cdot \boldsymbol{a}_n)|_{\Gamma_D \cup \Gamma_N} = 0 \,,$$

and in tangential direction

$$\left(\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{a}_n} \cdot \boldsymbol{a}_\alpha\right)\bigg|_{\Gamma_D \cup \Gamma_N} = 0 \qquad \text{for } \alpha = 1, 2 \,.$$

The velocity $\boldsymbol{w}$ is extended over the domain by solving the armonic operator $-\Delta \boldsymbol{w} = 0$ on Ω_f. Once $\boldsymbol{w}$ is known, the ALE map is defined as

$$\boldsymbol{x}_f(t) = \boldsymbol{x}_0 + \int_0^t \boldsymbol{w} \, d\tau \,, \tag{1.90}$$

that maps each point $\boldsymbol{x}_0$ of the reference configuration into the deformed configuration $\boldsymbol{x}_f(t)$.

Weak fluid-structure coupling

In the following, $(\cdot, \cdot)$ will denote the $L^2(\Omega(t))$ inner product, $(\cdot, \cdot)_{\Gamma_s}$ will denote the $L^2(\Gamma_s)$ inner product, and the bilinear form $a(\cdot, \cdot)$ is defined as

$$a(\boldsymbol{w}, \boldsymbol{v}) = \mu(\nabla \boldsymbol{w} + (\nabla \boldsymbol{w})^T, \nabla \boldsymbol{v}) \,.$$

Thus, we can write the weak formulation for the fluid problem as

$$\rho_f \left(\frac{\partial \boldsymbol{u}}{\partial t}\bigg|_{\mathcal{A}}, \boldsymbol{\phi}\right) + a(\boldsymbol{u}, \boldsymbol{\phi}) + \rho_f(((\boldsymbol{u} - \boldsymbol{w}) \cdot \nabla)\boldsymbol{u}, \boldsymbol{\phi}) - (p, \nabla \cdot \boldsymbol{\phi})$$
$$= \int_{\Gamma(t)} (\sigma_f \boldsymbol{a}_n) \cdot \boldsymbol{a}_n (\boldsymbol{\phi} \cdot \boldsymbol{a}_n) \, d\Gamma + \int_{\Gamma_N(t)} \boldsymbol{h} \cdot \boldsymbol{\phi} \, d\Gamma \,, \tag{1.91}$$
$$(q, \nabla \boldsymbol{u}) = 0 \,,$$

for all $\boldsymbol{\phi} \in \boldsymbol{H}^1(\Omega(t))$: $\boldsymbol{\phi}|_{\Gamma_D(t)} = 0$ and $(\boldsymbol{\phi} \cdot \boldsymbol{a}_\alpha)|_{\Gamma(t)} = 0$) and for all $q \in L^2(\Omega(t))$. Note that only the normal component of the stress on the boundaries $\Gamma(t)$ is highlighted.

With a similar approach, we can write the weak formulation for the shell equation (1.85) as

$$
\rho_s \varepsilon \Big(\frac{\partial^2 \eta}{\partial t^2}, \psi \Big)_{\Gamma_s} + (\beta^* \eta, \psi)_{\Gamma_s} + (P \nabla \eta, \nabla \psi)_{\Gamma_s} = (f_s, \psi)_{\Gamma_s} , \qquad (1.92)
$$

for all $\psi \in H^1(\Gamma_s)$.

The shell model allows us to reduce by one the dimension of the solid, so the structure equations can be reduced to a boundary condition on Γ_s for the solid problem. In particular, the shell boundary condition is imposed on the middle surface of the solid, as noted above.

Therefore, the two sub-systems (1.91) and (1.92) are coupled by imposing $\boldsymbol{\sigma}^f \cdot \boldsymbol{n} - f_s = 0$ on Γ_s. We define now the functional space $V^0 = \{ \boldsymbol{\phi} \in H^1(\Omega_f) : \boldsymbol{\phi}|_{\Gamma_{D,f}} = \mathbf{0} \}$, where $\Gamma_{D,f}$ are the boundaries of Ω_f where a Dirichlet conditions are imposed. In order to satisfy the continuity of the test functions $\boldsymbol{\phi} \cdot \boldsymbol{n} = \psi$ over the interface surface Γ_s in the coupled system, a new functional space is introduced as

$$
W^0 = \{ (\boldsymbol{\phi}, \psi) \in V^0 \times H^1(\Gamma_s) : \ \boldsymbol{\phi} \cdot \boldsymbol{n} = \psi \quad \text{over} \quad \Gamma_s \} . \qquad (1.93)
$$

We can now derive the weak form of the coupled final system by simplifying the two terms $\int_{\Gamma(t)} (\sigma_f \boldsymbol{a}_n) \cdot \boldsymbol{a}_n (\boldsymbol{\phi} \cdot \boldsymbol{a}_n) \, d\Gamma$ and $(f_s, \psi)_{\Gamma_s}$. Note that $\psi = (\boldsymbol{\phi} \cdot \boldsymbol{a}_n) \circ r(\boldsymbol{x})$. Thus, the coupled system reads

$$
\rho_f \Big(\frac{\partial \boldsymbol{u}}{\partial t} \Big|_{\mathcal{A}}, \ \boldsymbol{\phi} \Big) + \rho_f \big([(\boldsymbol{u} - \boldsymbol{w}) \cdot \nabla] \boldsymbol{u}, \ \boldsymbol{\phi} \big) + a(\boldsymbol{u}, \boldsymbol{\phi}) - (p, \nabla \boldsymbol{\phi})
$$
$$
\rho_s \varepsilon \Big(\frac{\partial^2 \boldsymbol{\eta}}{\partial t^2}, \psi \Big)_{\Gamma_s} + (\beta \boldsymbol{\eta}, \psi)_{\Gamma_s} + (P \nabla \boldsymbol{\eta}, \nabla \psi)_{\Gamma_s} = \int_{\Gamma_{N,t}} \boldsymbol{h} \cdot \boldsymbol{\phi} \, d\Gamma , \qquad (1.94)
$$
$$
(\nabla \cdot \boldsymbol{u}, \ q) = 0 ,
$$

for all $(\boldsymbol{\phi}, \psi) \in W^0$, $q \in L^2(\Omega_f)$. A finite element technique is used to obtain the discrete weak formulation of (1.94). With this approach the structural equation can be incorporated in the fluid equations as a Robin boundary condition.

1.4.3 Numerical modeling with FEM

We use a finite element technique to obtain the discrete weak formulation of (1.94). We threat explicitly the position of the fluid domain, and consider an implicit discretization of the coupling conditions. With this approach, the

structural equation can be incorporated in the fluid equations as a boundary condition (Robin scheme). The structural equation (1.92) can be put in strong form and discretized as

$$\rho_s \varepsilon \frac{\eta^{k+1} - 2\eta^k + \eta^{k-1}}{\Delta t^2} - \nabla(\boldsymbol{\sigma}\eta^{k+1}) + \beta\eta^{k+1} = f_s^{k+1} \ , \qquad (1.95)$$

where η^{k+1} is the unknown at the given iteration, and η^k and η^{k-1} are the solution at the last and second-last iteration, respectively. In the following, we will maintain this notation for all the unknowns. In addition, the interface conditions are discretized as

$$\frac{(\eta^{k+1} - \eta^k) \circ (\boldsymbol{x}_f^k)^{-1}}{\Delta t} = (\boldsymbol{u}^{k+1} \cdot \boldsymbol{a}_n) = u_n^{k+1} \qquad \text{on } \Gamma \ . \qquad (1.96)$$

While the movement of the domain is treated explicitly, the interface conditions (1.96) are treated implicitly.

Let now consider the terms of (1.94) related to the shell displacement, with the exception of the prestress term. The discretized problem can be transformed as

$$\int_{\Gamma_s} \left(\frac{\rho_s \varepsilon}{\Delta t^2}(\eta^{k+1} - 2\eta^k + \eta^{k-1}) + \beta\eta^{k+1} \right) \phi_n \circ \boldsymbol{x}_f^k \, d\gamma$$

$$= \int_{\Gamma_s} \left(\frac{\rho_s \varepsilon}{\Delta t} + \beta\Delta t \right) (u_n^{k+1} \circ \boldsymbol{x}_f^k)(\phi_n \circ \boldsymbol{x}_f^k) \, d\gamma$$

$$+ \int_{\Gamma_s} \left(\left(\frac{\rho_s \varepsilon}{\Delta t^2} + \beta \right) \eta^k + \frac{\rho_s \varepsilon}{\Delta t^2}\eta^{k-1} \right) (\phi_n \circ \boldsymbol{x}_f^k) \, d\gamma \qquad (1.97)$$

$$= \int_{\Gamma_s^k} \left(\frac{\rho_s \varepsilon}{\Delta t} + \beta\Delta t \right) u_n^{k+1} \phi_n J^k \, d\gamma$$

$$+ \int_{\Gamma_s^k} \left(\left(\frac{\rho_s \varepsilon}{\Delta t^2} + \beta \right) \eta^k + \frac{\rho_s \varepsilon}{\Delta t^2}\eta^{k-1} \right) \circ (\boldsymbol{x}_f^k)^{-1} \phi_n J^k \, d\gamma \, ,$$

where J^k is a jacobean taking into account the change of surface area going from the reference configuration Γ_s to the deformed one Γ_s^k.

The prestress term can be modeled using the two tensors

$$C = [\boldsymbol{a}_1, \boldsymbol{a}_2] \in \mathbb{R}^{3\times 2} \quad \text{and} \quad F^k = \nabla_{\boldsymbol{x}_0}\boldsymbol{x}_f^k \in \mathbb{R}^{3\times 3} \, ,$$

where $\boldsymbol{x}_0$ is the fixed system of coordinates. Now, we can express the prestress tensor in Cartesian coordinates as $\tilde{P} = CPC^T \in \mathbb{R}^{3\times 3}$. Therefore, the

prestress contribution in (1.94) can be discretize as

$$
\begin{aligned}
(P\nabla\boldsymbol{\eta}, \nabla\psi)_{\Gamma_s} \approx & \int_{\Gamma_s^k} \Delta t \left(F^k \tilde{P}(F^k)^T\right)\nabla u_n^{k+1} \cdot \nabla\phi_n \, d\gamma \\
& - \int_{\Gamma_s^k} \left(F^k \tilde{P}(F^k)^T\right)\nabla(\eta^k \circ (\boldsymbol{x}_n^f)^{-1}) \cdot \nabla\phi_n \, d\gamma
\end{aligned}
\tag{1.98}
$$

Now considering (1.97) and (1.98) we can consider the discrete problem as: find $\boldsymbol{u}_h^{k+1} \in V_{g,h}^t \subset V_g^t = \{\boldsymbol{v} \in H^1(\Omega(t)) : \boldsymbol{v}|_{\Gamma_D(t)} = \boldsymbol{g} \text{ and } (\boldsymbol{v} \cdot \boldsymbol{a}_\alpha)|_{\Gamma(t)} = 0\}$, and $p_h^{k+1} \in Q_h^k \subset L^2(\Omega(t))$ such that

$$
\begin{aligned}
& \frac{1}{\Delta t}(\boldsymbol{u}_h^{k+1}, \ \boldsymbol{\phi}_h)_k + a(\boldsymbol{u}_h^{k+1}, \boldsymbol{\phi}_h)_k + \left(((\boldsymbol{u}_h^k - \boldsymbol{w}_h^k) \cdot \nabla)\boldsymbol{u}_h^{k+1}, \ \boldsymbol{\phi}_h\right)_k \\
& - (p^{k+1}, \nabla\boldsymbol{\phi}_h)_k + \Delta t \int_{\Gamma_s^k} \left(\frac{\rho_s \varepsilon}{\Delta t^2} + \beta\right)u_{n,h}^{k+1}\phi_{n,h} + \widetilde{P}\nabla u_{n,h}^{k+1} \cdot \nabla\phi_{n,h} J^k \, d\gamma \\
& = \frac{1}{\Delta t}(\boldsymbol{u}_h^k, \ \boldsymbol{\phi}_h)_k - \int_{\Gamma_s^k} \left(\left(-\frac{\rho_s \varepsilon}{\Delta t^2} + \beta\right)\eta_h^k + \frac{\rho_s \varepsilon}{\Delta t^2}\eta^{k-1}\right) \circ (x_f^k)^{-1}\phi_{n,h}J^k \, d\gamma \\
& - \int_{\Gamma_s^k} \widetilde{P}\nabla(\eta^k \circ (x_f^k)^{-1}) \cdot \nabla\phi_{n,h}J^k \, d\gamma + \int_{\Gamma_{N,t}^k} \boldsymbol{h} \cdot \boldsymbol{\phi} \ J^k \, d\gamma \, , \\
& (\nabla \cdot \boldsymbol{u}_h^{k+1}, q_h)_k = 0 \ ,
\end{aligned}
\tag{1.99}
$$

where $\widetilde{P}$ is the prestress tensor in the reference configuration expressed in Cartesian coordinates, $\boldsymbol{\phi}_h \in V_h^k \subset W_0$, $q_h \in Q_h^k \subset Q^0$. Furthermore $u_{n,h}$ and $\phi_{n,h}$ are referred to the normal component of the vector to the outer surface and x_f^k is taken from equation (1.90) and maps each point of the simulated domain from the starting to the current configuration.

Once the velocity and the pressure field are computed, the displacement field can be obtained from

$$
\eta_h^{k+1} = \Delta t(u_{n,h}^{k+1} \circ x_f^k)|_{\Gamma_0} + \eta_h^k \, .
\tag{1.100}
$$

Then, the numerical problem is closed with the mesh motion, which is performed through a moving mesh algorithm based on a multigrid Arbitrary Lagrangian Eulerian method [57]. This technique allows to couple in an implicit way the interface conditions. As mentioned above, the points of the fluid domain are moved by solving an harmonic extension operator

$$
\begin{aligned}
& (\nabla\boldsymbol{w}_h^{k+1}, \nabla\boldsymbol{\psi}_h) = 0 \qquad \forall\boldsymbol{\psi}_h \in \boldsymbol{M}_h^0 \subseteq \boldsymbol{M}^0 \, , \\
& \boldsymbol{w}_h^{k+1}|_{\Gamma_s} = (\boldsymbol{u}_h^{k+1} \circ \boldsymbol{x}_f^k)|_{\Gamma_s} \, ,
\end{aligned}
\tag{1.101}
$$

with $\boldsymbol{M}^0 = \{\boldsymbol{\psi} \in \boldsymbol{H}^1(\Omega) \, : \, \boldsymbol{\psi}|_{\Gamma_s} = 0 \text{ and } (\boldsymbol{\psi} \cdot \boldsymbol{a}_n|_{\Gamma_D \cup \Gamma_N}) = 0\}$.

The numeric algorithm for the resolution of this problem can be summarized as in Algorithm 1.

Algorithm 1 Explicit Robin scheme for the numerical resolution of Koiter fluid-structure system.

---

1) Solve the system (1.99) on the domain Ω^k. Given $\boldsymbol{u}_h^k$, p_h^k and η_h^k we find $\boldsymbol{u}_h^{k+1}$ and p_h^{k+1}.
2) Compute the displacement field η_h^{k+1} with (1.100).
3) Compute the velocity $\boldsymbol{w}_h^{k+1}$ of the points of the fluid domain with (1.101).
4) Move the points of the fluid domain with

$$\boldsymbol{x}_f^{k+1}(\boldsymbol{x}_0) = \boldsymbol{x}_f^k(\boldsymbol{x}_0) + \Delta t \boldsymbol{w}_h^{k+1} \qquad \forall \boldsymbol{x}_0 \in \Omega \,.$$

1.4.4   Artificial boundaries for absorbing conditions

The presented Koiter model is often applied to hemodynamic simulations. In this framework, the fluid and the structure motion are most likely driven by pressure waves. Therefore, it is frequent to deal with numerical simulations involving only pressure inlet/outlet boundary conditions. In order to have a good representation of the pressure field in the simulated domain, it is necessary to introduce suitable outflow boundary conditions, in order to avoid spurious reflections of the pressure waves. In fact, even if the fluid is described by parabolic equations, FSI systems have some hyperbolic behavior. An approach based on artificial boundary conditions for truncated domains is adapted to prescribe outflow conditions able to absorb the pressure waves coming from the rest of the domain.

Boundary conditions for pressure and flux in bounded domains

As introduction of the mathematical and numerical problem, we consider the Navier-Stokes equation and we follow the method described in [58]. Let $\boldsymbol{b}$ be the extension of prescribed Dirichlet boundary values into the whole numerical domain Ω. Since we are considering a finite element discretization, this can be achieved by prescribing the appropriate nodal values along the boundary. As reported in Figure 1.2, we consider also Γ as the boundary walls (no-slip condition), S_1 as inlet boundary and S_2 and S_3 as outlet boundaries.
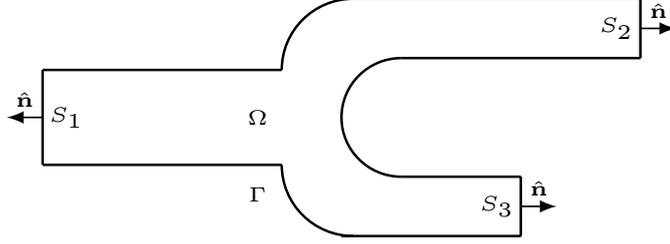
Figure 1.2: Reference duct with multiple truncated outlets (S_2 and S_3).

It is required that $\boldsymbol{u}(t) = \boldsymbol{b} + \boldsymbol{u}^*(t)$, where $\boldsymbol{u}^*(t) \in \boldsymbol{V}_1^*(\Omega) = \{\boldsymbol{\varphi} \in \boldsymbol{H}^1(\Omega) : \boldsymbol{\varphi}|_\Gamma = 0\}$ and $p(t) \in L^2(\Omega)$, $\forall t$. We have

$$\nu(\nabla\boldsymbol{u}, \nabla\boldsymbol{\varphi}) + (\boldsymbol{u}_t + \boldsymbol{u} \cdot \nabla\boldsymbol{u},\ \boldsymbol{\varphi}) - (p, \nabla \cdot \boldsymbol{\varphi}) = 0, \quad \forall\boldsymbol{\varphi} \in \boldsymbol{V}_1^*(\Omega) . \quad (1.102)$$

The pressure boundary conditions can be implicitly derived from (1.102). In particular, we can state that the mean pressure on each free section S_i is zero: $(|S_i|)^{-1} \int_{S_i} p\, ds = 0$. Now we formulate the problem more generally in terms of prescribed pressure drops. This can be achieved simply considering the equation (1.102) such that for any prescribed pressure π_i, the integral of the pressure gives $(|S_i|)^{-1} \int_{S_i} p\, ds = \pi_i(t)$.

Therefore, it is now required to find $\boldsymbol{u}(t)$ and $p(t)$ such that, for any prescribed π_i

$$\boldsymbol{u}_t + \boldsymbol{u} \cdot \nabla\boldsymbol{u} - \nu\Delta\boldsymbol{u} + \nabla p = 0, \qquad \nabla \cdot \boldsymbol{u} = 0, \qquad (1.103)$$

with

$$\boldsymbol{u}|_\Gamma = 0, \qquad \frac{1}{|S_i|} \int_{S_i} p\, ds = \pi_i(t) .$$

The numerical problem leads back to find the prescribed differences between the mean pressures across the various S_i (inlets and outlets). The variational problem reduces to find $\boldsymbol{u}(t) \in V_1^*(\Omega)$ and $p(t) \in L^2(\Omega)$ such that

$$\nu(\nabla\boldsymbol{u}, \nabla\boldsymbol{\varphi}) + (\boldsymbol{u}_t + \boldsymbol{u} \cdot \nabla\boldsymbol{u},\ \boldsymbol{\varphi}) - (p, \nabla \cdot \boldsymbol{\varphi}) = -\sum_j \pi_j(t) \int_{S_j} \boldsymbol{\varphi} \cdot \hat{\boldsymbol{n}}\, ds, \quad (1.104)$$

for all $\boldsymbol{\varphi} \in V_1^*(\Omega)$. Using the variational formulation (1.104) we can now derive a system of artificial boundary conditions. All the considerations made for the Navier-Stokes equations can now be applied to the studied Koiter-FSI problem, considering that we actually deal with the fluid equations together with an embedded Robin boundary condition. For this reason the extension

of the procedure (1.104) to the studied model is straightforward. Indeed one may wish to find the pressure drops that are required to achieve a desired net flux through each of various ducts. Since in this work we use prescribed inlet pressure, in our case this search is limited to the definition of the reference pressures at the outlets π_j.

The absorbing boundary conditions

We prescribe an absorbing boundary condition by coupling the 3D model with a 1D reduced one [59], in order to obtain a consistent value of pressure π_j to be imposed at the outlets of our domain. The guess of pressure field at the outlet is obtained from the simplified 1D model. For this purpose we consider a cylinder whose length is L. The simplified 1D model can be obtained by integrating at each time t the Navier–Stokes equations over each section S normal to the axis of the cylinder. For each $t > 0$ and $0 < z < L$ the 1D model reads

$$\begin{aligned}
&\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial z} = 0, \\
&\frac{\partial Q}{\partial t} + \frac{\partial}{\partial z}\Big(\alpha \frac{Q^2}{A}\Big) + \frac{A}{\rho_f}\frac{\partial P}{\partial z} + K_R \frac{Q}{A},
\end{aligned} \tag{1.105}$$

where Q is the flow rate through S, A the area of S, P the mean pressure over S, K_R a resistance parameter which accounts for the fluid viscosity and α accounts for the shape of the velocity profile over S [60]. For the closure of system (1.105), a third equation is provided through a pure algebraic wall model, relating the radial displacement to the mean pressure in a section $P = \frac{\beta(\sqrt{A}-\sqrt{A^0})}{\pi}$, where A^0 is the area of the surface S at $t = 0$, h_s the wall thickness, E the Young modulus of the solid wall, ν the Poisson coefficient and β is given by (1.86).

The system now turns out to be hyperbolic, and it possesses two distinct eigenvalues. The absorbing outflow boundary condition is derived by imposing that the characteristic variable entering the 3D computational domain be zero, meaning that no information is entering. In particular, we impose

$$W_2|_{\Gamma_{out}^i} = \left[\frac{Q}{A} \pm \frac{2\sqrt{2}}{\sqrt{\rho_f}}\Big(\sqrt{P + \beta\sqrt{A^0}} - \sqrt{\beta\sqrt{A^0}}\Big)\right]\Bigg|_{\Gamma_{out}^i} = 0, \tag{1.106}$$

obtaining

$$P|_{\Gamma_{out}^i} = \pi_j(t) = \left[\Big(\frac{\sqrt{\rho_f}}{2\sqrt{2}}\frac{Q}{A} + \sqrt{\beta\sqrt{A^0}}\Big)^2 - \beta\sqrt{A^0}\right]\Bigg|_{\Gamma_{out}^i}. \tag{1.107}$$

By replacing $\pi_j(t)$ obtained in (1.107) into (1.104) we can obtain suitable Neumann inhomogeneous outflow boundary conditions that allow us to treat multidimensional phenomena along the studied domain.

1.5 Numerical results

In this section the numerical results of the presented fluid-structure interaction simulations are presented. The algorithm for the numerical implementation of the monolithic FSI model presented in section 1.3 into the finite element code FEMuS has been presented in many works involving both multiphysics studies (e.g. see [61]) and optimal control applications (e.g. see [62, 63, 6]). Therefore, in this section we focus on the results of the Koiter model implemented in this work, together with the absorbing boundary conditions.

1.5.1 Monolithic fluid-structure benchmark results

The numerical benchmarking of the presented monolithic FSI solver has been already presented in some previous works [6]. Here some of the most significant results presented in the cited work are briefly reported. Then, in the next section, the numerical results on the Koiter shell fluid structure model developed in this work are reported. Numerical benchmark for both standalone CFD (Navier Stokes) and CSM (Computational Structural Mechanics) benchmarks have been considered. Thus, the fluid-solid mutual interaction is neglected. All the tested cases have shown good results in agreement with the benchmark solutions. Then, the benchmark results for the complete FSI problem have been developed.

Table 1.1: Parameter setting for the FSI1 benchmark.

Parameter	Symbol	Measure unit	Value
Solid density	ρ_s	$[10^3 \frac{kg}{m^3}]$	1
Poisson coefficient	ν_s	-	0.4
Shear modulus	μ_s	$[10^6 \frac{kg}{ms^2}]$	0.5
Fluid density	ρ_f	$[10^3 \frac{kg}{m^3}]$	1
Fluid viscosity	ν_f	$[10^{-3} \frac{m^2}{s}]$	1
Average inlet velocity	$\bar{U}$	$[\frac{m}{s}]$	0.2

Table 1.2: Results for the Turek FSI1 benchmark.

Level	Nel	Ndof	Dx(A)(e-05) $[m]$	Dy(A)(e-04) $[m]$	Drag	Lift
1	1870	17330	2.15828	8.34873	14.4034	0.750155
2	9350	68320	2.16161	8.28078	14.3377	0.757567
3	39270	271280	2.16367	8.23453	14.3074	0.761073
Ref			2.27049	8.20877	14.2943	0.763746

We report only the benchmark on the mutual interaction between solid and fluid, called Turek FSI1 test [1]. In Table 1.1 we show the values of the parameter used in this test. In Table 1.2 the results obtained with different spatial resolutions are reported and compared with the reference values. The results converge to the values given by the benchmark authors, when the mesh resolution is increased. The presented and validated numerical code for FSI simulations can be found in the Github page [18]. Since the implemented model has been validated with the Turek benchmark, in the following it will be used as a reference case to validate also the implemented Koiter model.

1.5.2   Koiter FSI model numerical benchmarks

Since the presented fluid-structure model based on Koiter shell equations is still not widespread, in literature there are only a few works on a benchmarking of it. In this section, we will refer to the benchmarks presented in [64]. In particular, we first consider a simple numerical case where an analytical solution is available. We refer to it as *Koiter benchmark 1*. Another benchmark based on a comparison between the Koiter model and the monolithic fluid-structure is then presented. We refer to it as *Koiter benchmark 2*. Since the monolithic fluid-structure model has been widely tested, as reported in the previous section, we can use it as a reference result for the new presented model. However, since the two models are different, we expect a similar (but not identical) behavior of the two models under the same parameter setting.

Koiter benchmark 1

We consider now a simple FSI problem that has an exact solution, and can be used for the testing of the presented fluid-structure shell model. We consider a fluid flowing through a cylindrical channel, with the inlet at the bottom and outlet at the top, as can be seen in Figure 1.3. On the outer wall the

Robin boundary condition for Koiter shell is imposed. In the Figure on the left, the mesh used for this benchmark is reported.
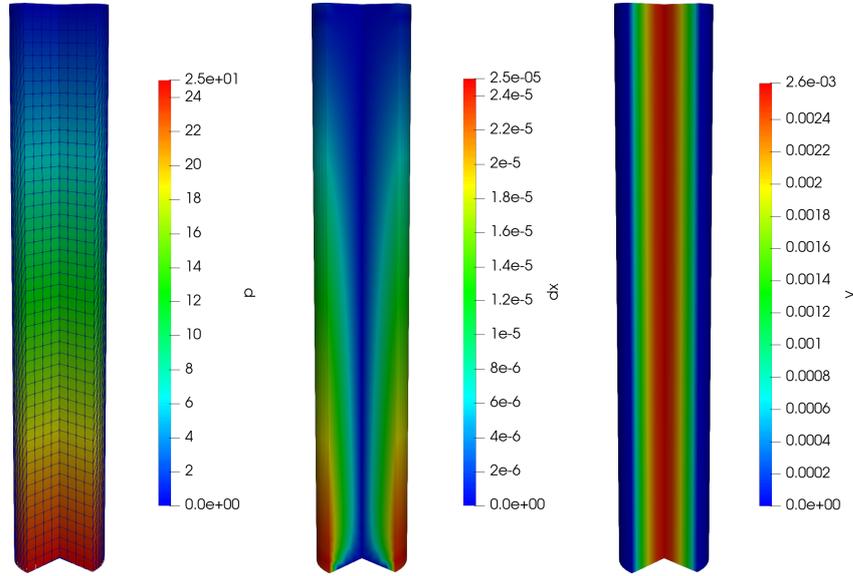
Figure 1.3: Pressure p (left), displacement dx (center) and velocity v (right) fields in the simulated channel. On the left is also reported the mesh.

Table 1.3: Parameter setting for *Koiter benchmark 1*.

Parameter	Symbol	Measure unit	Value
Cylinder radius	R	$[m]$	$5 \cdot 10^{-3}$
Cylinder length	L	$[m]$	$6 \cdot 10^{-2}$
Solid density	ρ_s	$[10^3 \frac{kg}{m^3}]$	1.1
Poisson coefficient	ν_s	-	0.1
Elastic modulus	E	$[kPa]$	125
Wall thickness	ε	$[m]$	$2 \cdot 10^{-4}$
Fluid density	ρ_f	$[10^3 \frac{kg}{m^3}]$	1
Fluid viscosity	ν_f	$[\frac{m^2}{s}]$	$1 \cdot 10^{-3}$

The shell equation to be solved in this case (for cylindrical geometries) is (1.87), where the prestress term is neglected. Under this hypothesis, the problem turns to be linear, and can be exactly solved. By imposing a constant inlet pressure p_{in}, it is possible to obtain the analytical solution of the pressure and the displacement fields for stationary solutions. All the variables involved

in this benchmark are reported in Table 1.3. Under the presented parameter setting we have for cylindrical geometries

$$\beta = \frac{\varepsilon E}{(1 - \nu_s)R^2} = 1010 kPa/m \,,$$

as reported in (1.86). By setting a time step of $0.0005s$, an inlet pressure of $p_{in} = 25Pa$ and an outlet pressure of $p_{out} = 0Pa$, after a time $t = 0.25s$ the steady state is reached. In Figure 1.3, the pressure, displacement and velocity fields at the steady state are qualitatively reported.
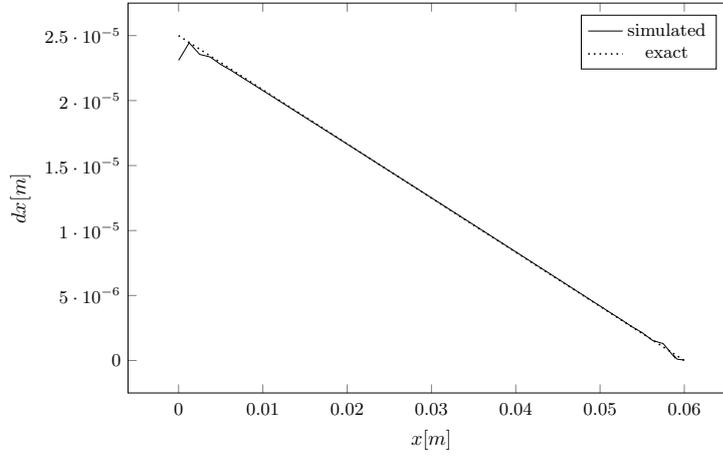
Figure 1.4: Comparison of the displacement field dx between the simulated case and the reference one. The displacement field is reported along the line between the points $(0.005, 0, 0)$ and $(0.005, 0, 0.6)$.

As mentioned in [64], the fluid pressure is linear within the channel as

$$p_e(r, z) = p_e(z) = \frac{p_{out}z + p_{in}(L - z)}{L} \quad z \in [0, L], \, , \forall r \in [0, R] \,. \qquad (1.108)$$

Thus the exact radial displacement of the structure is simply given by

$$\eta_e(z) = \frac{p_e(z)}{\beta} \,. \qquad (1.109)$$

The comparison between the displacement field simulated with the implemented algorithm and the exact displacement is reported in Figure 1.4. It can be noted that there is a good agreement between the expected (1.109) and the simulated values of the considered field. However, some discrepancies between the exact and the simulated solutions can be found at the extremes
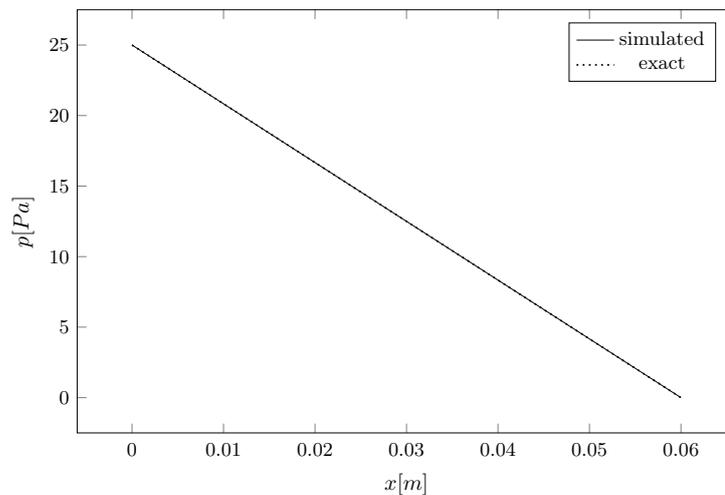
Figure 1.5: Comparison of the pressure field p between the simulated case and the reference one, along the cylinder axis.

of the cylinders (at the inlet and outlet). This is due to the boundary conditions on the displacement field imposed on the corner of the domain. In future works, the boundary conditions on the edges will be analyzed in detail and improved.

At the same time, in Figure 1.5 the comparison between the exact (1.108) and the simulated pressure fields along the cylinder axis is reported. It can be seen that there is total agreement between the simulated pressure and the exact one.

Koiter benchmark 2

The second benchmark presented in this section, as well as the first one, is taken from the same work cited above. The idea of this benchmark is to make a comparison between the two different fluid-structure models presented in sections above. In the cited work a full, nonlinear FSI problem with the structure consisting of two layers has been solved. In particular, one layer with thickness h_1 was simulated with the Koiter FSI model, and one layer with thickness h_2 has been simulated with a full three dimensional fluid structure model. The combined thickness of the composite structure was considered constant, and the benchmark consists of finding similar results by changing h_1 and h_2.

In this section, we do not consider a multi-layer case, and the compari-
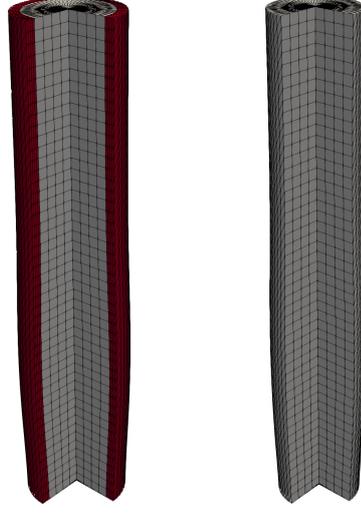
Figure 1.6: The two meshes used for the numerical simulation of Koiter benchmark 2. The mesh for the monolithic FSI model (left, in red the solid domain) and the mesh for the Koiter FSI model (right) are shown.

son is carried out between the Koiter FSI model and the monolithic model presented above. Since the monolithic model satisfies the Turek benchmark, as reported in section 1.5.1, we can use it as a reference for the testing of the new algorithm. The test has been carried out using the same parameters of the Koiter benchmark 1 reported in Table 1.3, with the only exception of the elastic modulus and the wall thickness. In fact, in this case $E = 15625 Pa$ and $\varepsilon = 1.2 \cdot 10^{-3} m$. Using the same simplified relation for the β calculus, we obtain $\beta \simeq 757 kPa/m$.

In order to test the temporal evolution of the two implemented algorithms, a variable inlet pressure value is considered

$$p_{in} = \begin{cases} \frac{p_{max}}{2} \left(1 - \cos \left(\frac{2\pi t}{t_{max}} \right) \right) & \text{if } t < t_{max}, \\ 0 & \text{if } t \geq t_{max}, \end{cases}$$

where $p_{max} = 1.333 kPa$ and $t_{max} = 0.15 s$. With all the presented hypotheses, the two different tests have been carried out on the meshes presented in Figure 1.6, with a time step $\Delta t = 5 \cdot 10^{-4} s$. Note that the fluid meshes for both the monolithic and Koiter FSI models are coincident, while the mesh for the monolithic model is equipped also with the solid mesh. In Figure, the fluid enters from the bottom and exits from the top. On the outer wall (external surface of the cylinders) the structure conditions are imposed.
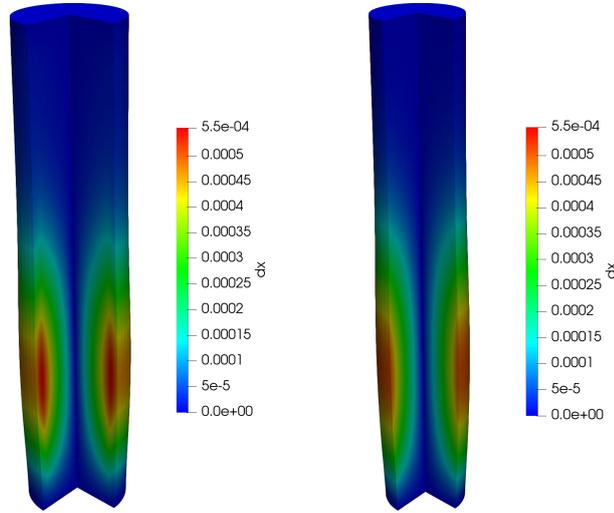
Figure 1.7: Displacements dx fields for both the monolithic model (left) and the Koiter model (right).

In Figure 1.7 the qualitatative behavior of the dx field is reported for both the monolithic and the Koiter model for $t = 0.04s$. It can be seen that the two cases have similar displacement fields. Also, the velocity and pressure fields present similar behavior, but are not reported in this framework.

In Figure 1.8, the quantitative comparison of the displacement of the structure, calculated along the line between the two points $(0.005, 0, 0)$ and $(0.005, 0, 0.6)$, is reported. We recall that the problem is symmetric with respect to the cylinder axis, so the reported graphs are valid for every line parallel to the cylinder axis and located on the external surface of the unde-formed cylinder. Note that the two models have a similar behavior in time, for all the reported time steps. In particular, the two models differ to each other for $t = 0.02s$ (top-left in Figure) and for $t = 0.04s$ (top-right). After the initial transition, the two solutions seem to converge to a similar solution, as it can be seen in the bottom part of Figure 1.8, for $t = 0.06s$ and $t = 0.08s$.

In order to show the evolution in time of the difference between the two solutions, the L^2 norm of the displacement dx is considered. In particular, we have

$$\|dx\| = \sqrt{\int_\Omega dx^2 \, d\boldsymbol{x}}\,,$$

where dx is derive from the extension of the displacement η inside the domain through the Laplacian operator introduced above. With this notation we
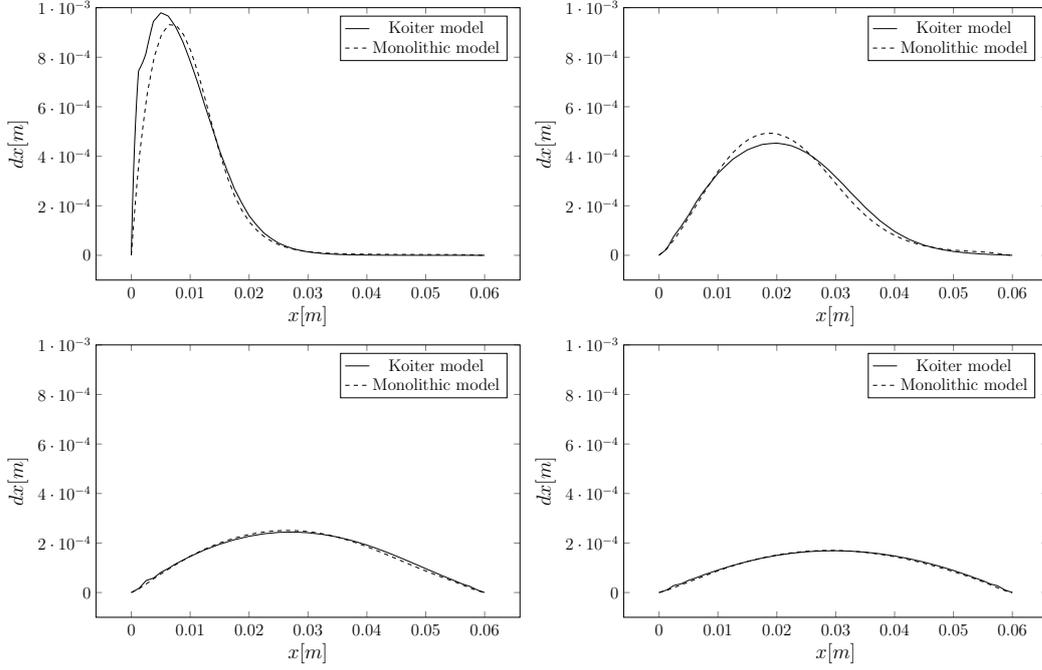
Figure 1.8: Comparison between Koiter (continuous line) and the monolithic model (dashed line) displacement, for $t = 0.02s$ (top left), $t = 0.04s$ (top right), $t = 0.06s$ (bottom left) and $t = 0.08s$ (bottom right).
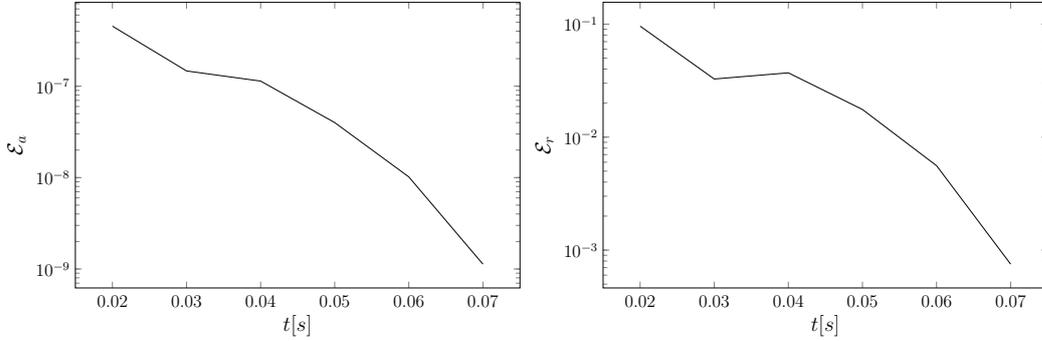
Figure 1.9: Temporal accuracy of the model through absolute (left) and relative (right) errors between the norm of dx in monolithic and Koiter models.

study the absolute norm error between the two models defined as

$$\mathcal{E}_a = \left| \|dx\|_m - \|dx\|_k \right|,$$

where $\| \cdot \|_m$ is the L^2 norm on Ω in the monolithic case, and $\| \cdot \|_k$ is the L^2

norm on Ω in the Koiter case. We also study the relative error defined as

$$\mathcal{E}_r = \frac{\mathcal{E}_a}{\|dx\|_m} \; .$$

In figure 1.9 we report the absolute error (on the left, in log-scale) and the relative error (on the right) between the two solutions. The numerical Koiter solution tends to the numerical monolithic solution. Therefore, we can assert that the implemented Koiter model is consistent with the solution obtained with the monolithic model, which has been fully tested in previous works.

We can finally assert that the implemented code shows good results according to the presented simple benchmarks. Interested reader in more benchmarks on the Koiter shell equations applied to fluid-structure interaction, and in benchmarks for composite structures can see [64].

1.5.3 Grid convergence

Another non-trivial issue, especially when dealing with moving meshes, is the grid convergence of the implemented algorithm. In particular, in this section we study the grid convergence of the full Koiter model, with the linear term and the prestress term. In order to show it, a multigrid technique is used to study the behavior of the solution when the grid is refined. Interested reader in multigrid technique can see [65, 57].

Three tests on different cases will be carried out, in order to test the convergence on different geometries and boundary conditions. In particular, we consider a plane channel, a two-dimensional airbag-type rectangle and a cylindrical channel. The main parameter used to test the convergence is the displacement $dx = \eta$.

The plane channel test

A simple plane channel of size $0.1m \times 0.3m$ is first considered, as reported in Figure 1.10. It is meshed with a simple 2×2 grid, with quadratic elements, for a total of 25 degree of freedom. By considering the undeformed plane channel with each side parallel to the Cartesian axes, we impose a pressure inlet condition on the bottom side Γ_1 (constant pressure in time, $p_{in} = 6000Pa$), a no-slip condition on the left wall Γ_2, an outflow condition on the top side Γ_4, and the Koiter boundary conditions on the right wall Γ_3 (see Figure 1.10). The values of the parameters for all the grid convergence tests are reported in Table 1.4.
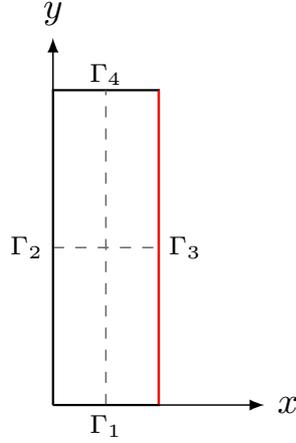
Figure 1.10: Domain for all the convergence tests. For the cylindrical channel, a rotation around the y-axis of the presented geometry is considered.

Table 1.4: Parameter setting for all the grid convergence tests.

Parameter	Symbol	Measure unit	Value
Solid density	ρ_s	$[10^3 \frac{kg}{m^3}]$	1
β coefficient	β	$[kPa/m]$	200
Wall thickness	ε	$[m]$	$7.5 \cdot 10^{-3}$
Fluid density	ρ_f	$[10^3 \frac{kg}{m^3}]$	1
Fluid viscosity	ν_f	$[\frac{m^2}{s}]$	$1 \cdot 10^{-3}$

We report the stationary solution for all the tested cases. We consider now a time step of $\Delta t = 0.1s$. After 50 iterations, a steady state is found. In Figure 1.11, the solution of the displacement field along the line between the points $(0.1, 0)$ and $(0.1, 0.3)$ (red line in Figure 1.10) is reported. In particular, the solution is reported for various mesh refinements, ranging from 1 level (the introduced 2×2 mesh) to 4 level. Since the solutions with 3 and 4 levels are practically the same, we can conclude that the solution converges to a certain field.

We have also considered the L^2-norm of the solution calculated over the entire domain, similarly to the last section. We report the results in Table 1.5. Since the exact solution is unknown, we use a common approach to determine the convergence rate p of the solution, given by

$$p = \log_2 \frac{\|dx\|_{L-1} - \|dx\|_{L-2}}{\|dx\|_L - \|dx\|_{L-1}} , \qquad (1.110)$$
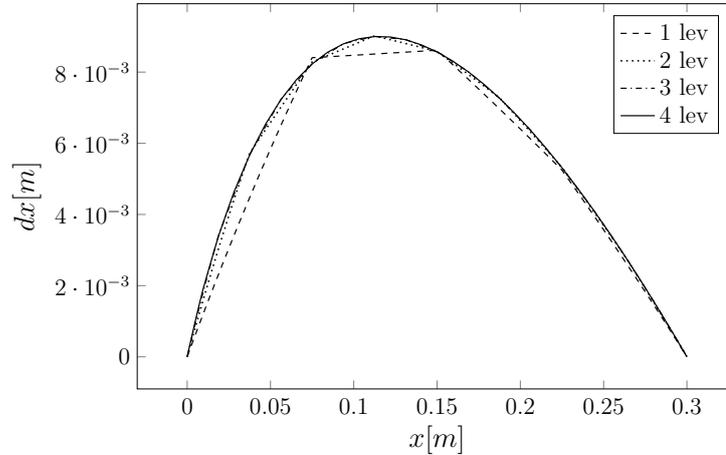
Figure 1.11: Comparison between the displacement field solution for different grid refinements.

Table 1.5: Convergence of the L^2-norm of the displacement field for all the tested cases. In particular, the convergence p for unknown exact solution is reported.

Level L	N_{elements}	$\|dx\|$	$\big\|\|dx\|_L - \|dx\|_{L-1}\big\|$	p
1	4	$7.2024 \cdot 10^{-4}$	$-$	$-$
2	16	$7.1024 \cdot 10^{-4}$	$9.9952 \cdot 10^{-6}$	$-$
3	64	$7.0674 \cdot 10^{-4}$	$3.5039 \cdot 10^{-6}$	1.51
4	256	$7.0578 \cdot 10^{-4}$	$9.5290 \cdot 10^{-7}$	1.88
5	1024	$7.0551 \cdot 10^{-4}$	$2.7350 \cdot 10^{-7}$	1.80

where L is the level of refinement. We report the calculated values of p in Table 1.5, together with the absolute error between the refined level L and $L-1$. Also, in this case the code shows good convergence properties. Note that in this section we are not interested in the accuracy of the presented results, but only on the convergence of the solution to a value.

The airbag test

We consider now the same geometry of the previous case (see Figure 1.10), with different boundary conditions. In particular, a constant pressure is imposed on the boundary Γ_2 of $p_{in} = 6000Pa$, a no-slip condition is imposed on Γ_4 and Γ_1, and the Koiter shell condition is imposed on the wall Γ_3. All the parameters are considered equal to the set introduced in Table 1.4.
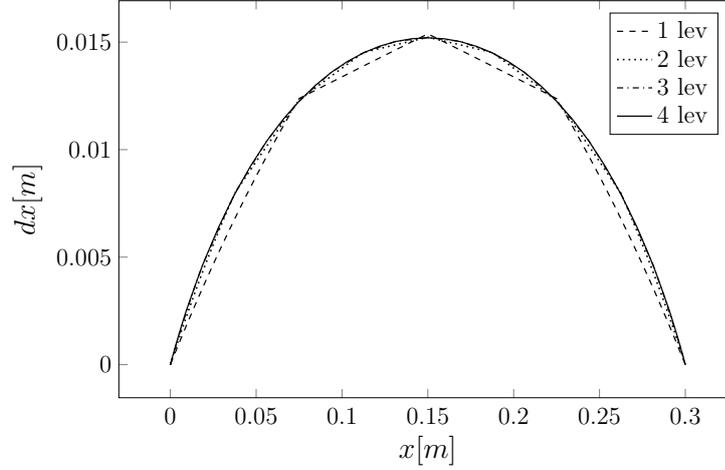
Figure 1.12: Comparison between the displacement solution for different grid refinements in the airbag test.

In Figure 1.12, the comparison between the solution of the displacement dx for different refinement levels is reported. Again, the solution is plotted along the red line in Figure 1.10. Note that the solution is symmetric, as expected, and seems to converge with the grid.

Table 1.6: Convergence of the L^2-norm of the displacement field for the airbag test. The convergence p for unknown exact solution and the absolute error between consecutive solutions are reported.

Level L	N_{elements}	$\|dx\|$	$\left\| \|dx\|_L - \|dx\|_{L-1} \right\|$	p
1	4	$1.43798 \cdot 10^{-3}$	—	—
2	16	$1.42065 \cdot 10^{-3}$	$1.73299 \cdot 10^5$	—
3	64	$1.41449 \cdot 10^{-3}$	$6.15841 \cdot 10^6$	1.49
4	256	$1.41279 \cdot 10^{-3}$	$1.69774 \cdot 10^6$	1.85
5	1024	$1.41213 \cdot 10^{-3}$	$6.58422 \cdot 10^7$	1.36

In Table 1.6, the L^2-norm of the dx solution on the entire domain is again reported. The absolute error and the convergence rate p, calculated as in equation (1.110), are also reported. As in the previous case, it can be noted that the solution seems to converge with the grid, with a $1 < p < 2$ in all the tested cases.

The cylindrical channel test

We finally consider a simple cylindrical geometry in order to show the convergence properties of the three-dimensional numerical problem. The physical quantities involved in this test are the same used for the previous tests. The cylinder is obtained by a rotation of the geometry 1.10 around the y-axis. We impose a constant pressure on the bottom surface $p_{in} = 6000 Pa$, an outflow condition on the top surface, and the Koiter shell boundary equation on the outer wall.
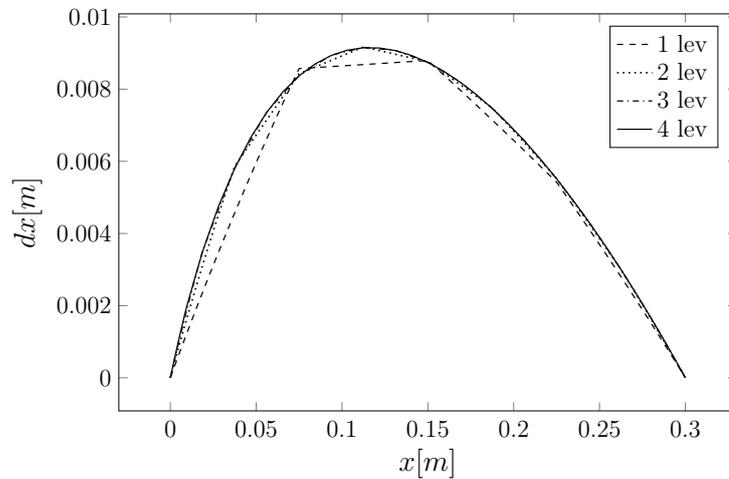
Figure 1.13: Comparison between the displacement solution for different grid refinements in the cylinder test.

In Figure 1.13, the comparison between the solution dx along a line on the outer surface of the cylinder and parallel to the axis of rotation of the cylinder is shown for different mesh refinement. It can be noted that, again, the solution tends to settle to a certain profile, in agreement with the convergence requirements.

1.5.4 Absorbing boundary condition test

In this section we introduce a test in order to show the robustness of the algorithm developed for the absorbing boundary conditions presented in section 1.4.4. In particular, this test is inspired by the test presented in [5]. The considered geometry is a cylindrical channel obtained from the rotation of the rectangle of dimension $0.05m \times 1m$ around the y-axis (see Figure 1.14).
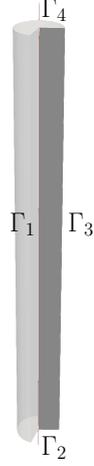
Figure 1.14: Domain for all the absorbing tests. The rectangular geometry is rotated around the y-axis.

The rectangle has its bottom-left corner on the axis origin, and is meshed with a 16×100 uniform rectangular grid. In all the tests presented in the following, we consider the parameters presented in Table 1.7, with a temporal discretization of $\Delta t = 0.001s$. Each presented test is characterized

Table 1.7: Parameter setting for all the tests on the absorbing boundary conditions.

Parameter	Symbol	Measure unit	Value
Solid density	ρ_s	$[10^3 \frac{kg}{m^3}]$	1
β coefficient	β	$[kPa/m]$	2000
Wall thickness	ε	$[m]$	$7.5 \cdot 10^{-3}$
Fluid density	ρ_f	$[10^3 \frac{kg}{m^3}]$	1
Fluid viscosity	ν_f	$[\frac{m^2}{s}]$	$1 \cdot 10^{-3}$

by a different value of the inlet pressure (imposed on the bottom wall Γ_1). We also impose an outflow boundary condition on the top wall Γ_4, and we compare it with the same case with absorbing outflow boundary conditions. The Koiter shell equation boundary condition is also imposed on the right wall Γ_3. Finally, a condition consistent with the studied axisymmetric case is imposed on the left wall Γ_1. Note that in order to carry out these tests based on the axisymmetric boundary condition, it is important to derive the correct formulations for all the mathematical operators involved in the PDE

to be solved.

We now report two cases to test the absorbing boundary conditions, based on a step function and on a sinusoidal inlet pressure in time.

Step function inlet pressure

In this section, we consider the presented system with a step function in time inlet pressure and compare the response with classical outflow boundary conditions and absorbing outflow boundary conditions. The imposed pressure reads

$$p_{in} = \begin{cases} 5000\,Pa & \text{if } t < t_{max}\,, \\ 0\,Pa & \text{if } t \geq t_{max}\,, \end{cases}$$

where $t_{max} = 0.03s$. In particular, we expect that the pressure step function induces a pressure wave in the studied domain. We also expect that, after transiting inside the domain, the pressure wave leaves it and the mean pressure inside the domain goes to zero.

In Figure 1.15, the transition of the pressure along the symmetry axis of the cylinder is reported. In the first time steps after the transition from $p_{in} = 5000Pa$ to $p_{in} = 0Pa$, the pressure fields in the two cases (with outflow boundary condition and with absorbing boundary condition) are the same, as can be seen in Figure for $t = 0.03s$ and $t = 0.06s$. For $t = 0.09s$, the outflow boundary condition starts to influence the exit of the pressure wave from the domain, therefore the two pressure profile are different. Note that while the pressure at the outlet is prescribed to be zero in the outflow case, in the absorbing case the pressure is equal to the pressure expected from the equivalent one-dimensional system, as introduced in section 1.4.4. From $t = 0.12s$ spurious waves can be noted in the outflow case. The spurious wave is originated by the non-physical imposition of a constant pressure at the outlet, which acts as a reflector of waves.

The error made by using a classical outflow approach can be quantified through the L^2-norm of the pressure over the entire domain. In fact, after the step of the pressure at $t = t_{max}$, the physics of the problem requires that the pressure wave leaves the domain and the pressure in the domain goes to zero. The L^2-norm of the pressure calculated on the domain for different time steps and for both the outflow and the absorbing case is reported in Table 1.8. Note that, while the value of the norm of the pressure in the outflow case has an unstable behavior in time caused by the reflections of the pressure
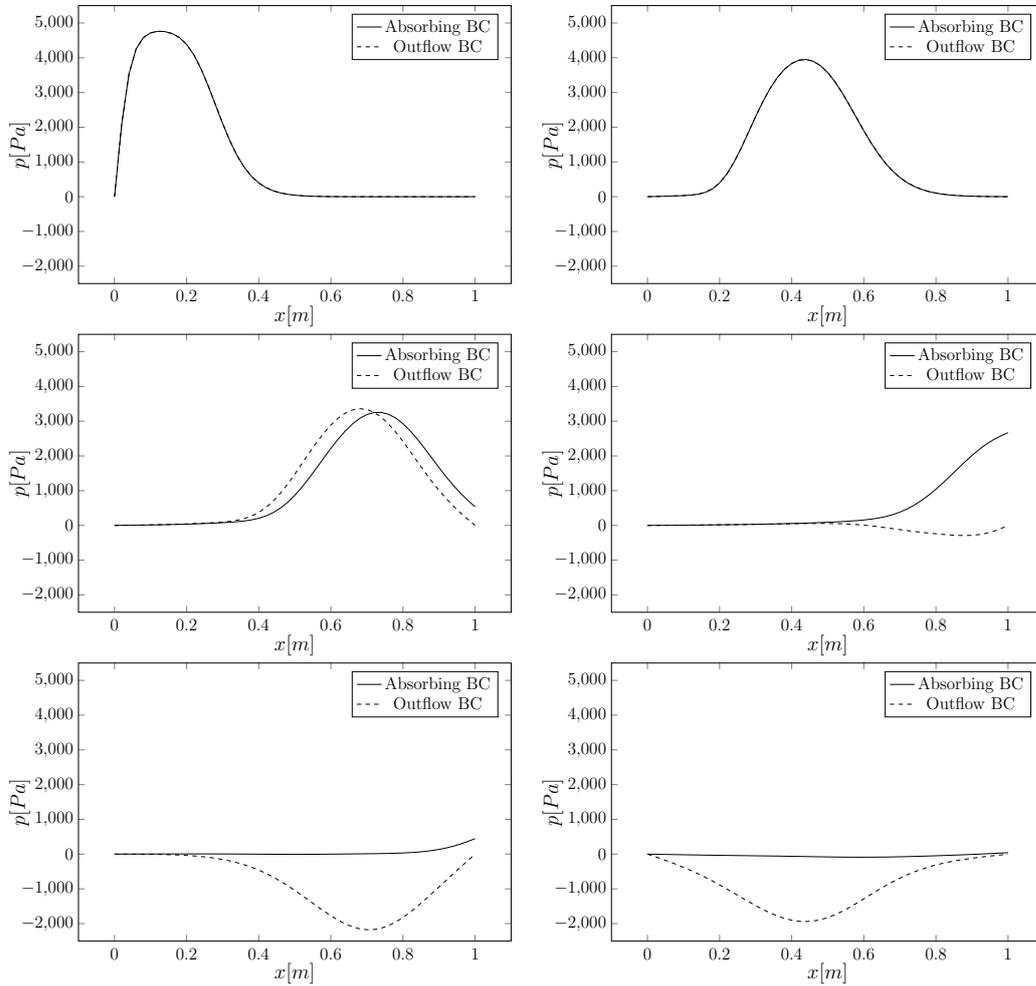
Figure 1.15: Pressure field along the cylinder axis for different time steps, starting from the time t_{max}. The pressures are reported for $t = 0.03s$ (top left), $t = 0.06s$ (top right), $t = 0.09s$ (center left), $t = 0.12s$ (center right), $t = 0.15s$ (bottom left) and $t = 0.18s$ (bottom right).

wave, the norm in the absorbing case goes to zero.

Therefore, we can conclude that the absorbing case is a better representation of the simulated physical system when dealing with impulsive inlet pressure. However, in Table 1.8 we can see that the norm in the absorbing case doesn't go identically to zero once the pressure wave exits the domain. This is due to very small spurious reflections related to the fact that the one-dimensional system used to predict the p value at the outlet of the cylinder is not accurate as a full three dimensional solver.

Table 1.8: L^2 norm of the pressure over the entire domain at various time step for both the outflow (p_{out}) and the absorbing case (p_{abs}).

Time $[s]$	$\|p\|_{out}$	$\|p\|_{abs}$
0.2	115.36	15.50
0.3	158.94	4.34
0.4	116.62	2.03
0.5	84.64	0.09

Sinusoidal inlet pressure

We consider now a sinusoidal inlet pressure to simulate a physical system close to the hemodynamic simulations. In fact, the blood vessels are often subject to sinusoidal-like inlet pressures boundary conditions. Simulate such cases without accounting for the absorbing conditions leads to serious errors in the simulations.
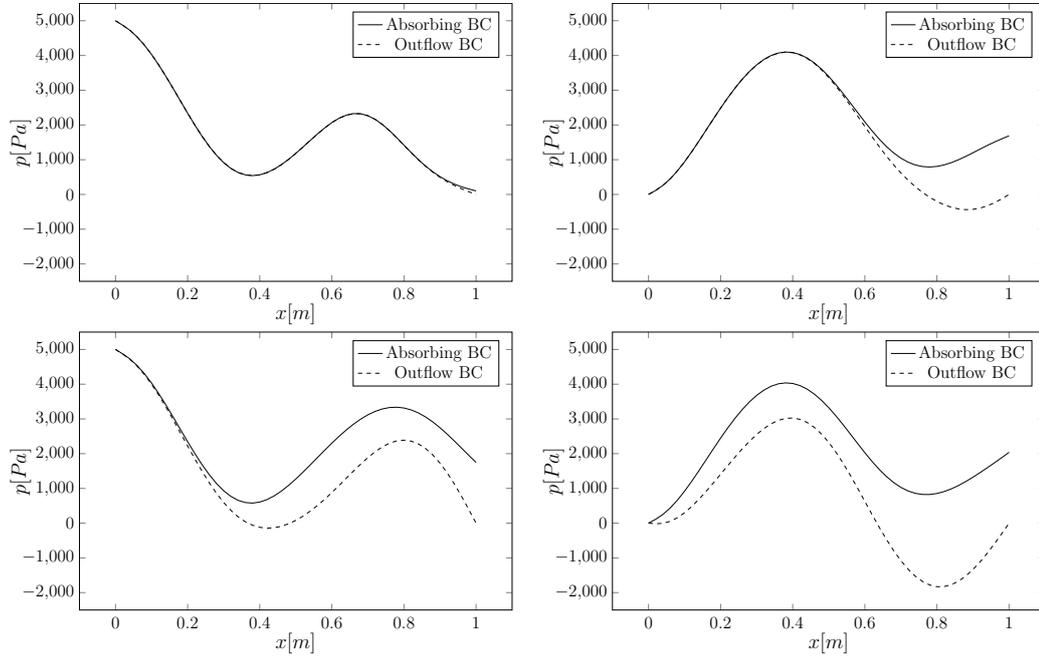
Figure 1.16: Comparison of the pressure field along the cylinder axis for different time steps between the same simulation with different outflow boundary conditions. Every image is out of phase of π with its subsequent.

We consider then a sinusoidal inlet pressure defined as

$$p_{in} = \frac{p_{max}}{2}\left(1 + \cdot \cos\left(\frac{\pi t}{T}\right)\right),$$

where $p_{max} = 5000Pa$ and $T = 0.04s$. We use a time step of $\Delta t = 0.001s$.
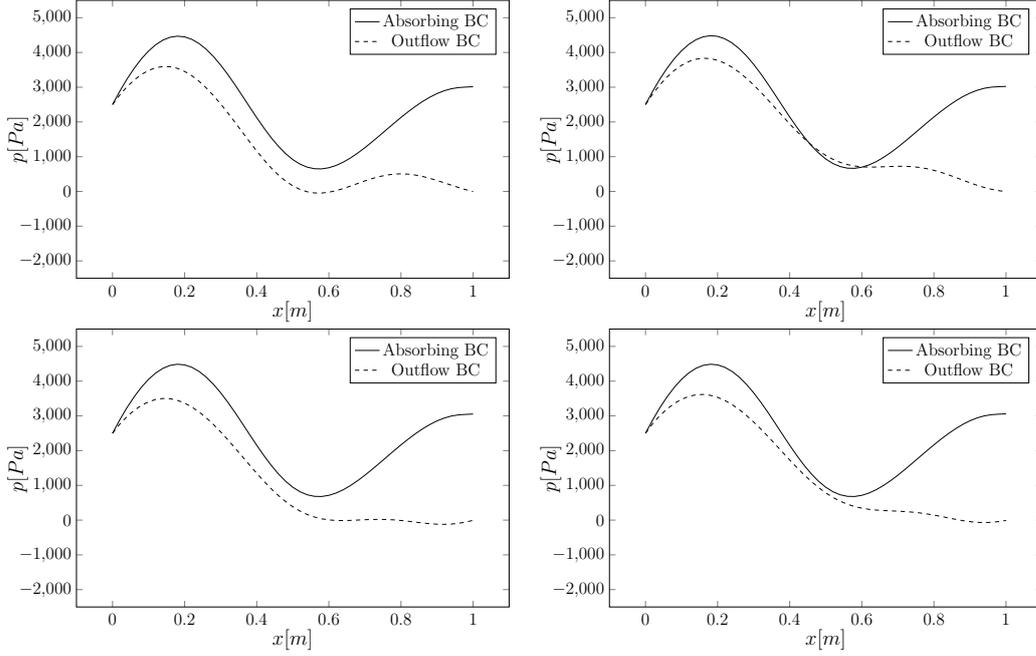
Figure 1.17: Comparison of the pressure field along the cylinder axis for the in-phase time steps $t = 0.26s$ (top left), $t = 0.34s$ (top right), $t = 0.42s$ (bottom left), $t = 0.50s$ (bottom right) between the same simulation with different outflow boundary conditions.

In Figure 1.16 we report the pressure fields along the cylinder axis at the start of the numerical simulation, in order to evaluate the first interaction of the induced pressure waves with the outflow boundaries. In particular, we report the pressure field for $t = 0.08s$ (top left), $t = 0.12s$ (top right), $t = 0.16s$ (bottom left), $t = 0.20s$ (bottom right). We remark that, initially, the solutions coincide ($t = 0.08s$) then, when the pressure waves hit the outflow boundary, the classical outflow boundary conditions affect the solution ($t = 0.12s$ and $t = 0.16s$) and in particular for $t = 0.20s$ a spurious negative wave is reflected and interact with the upcoming newly generated wave. At the same time, the absorbing boundary condition works properly and no spurious reflections can be noticed in Figure for all the reported time

steps. Moreover, since the four reported steps are out of phase of π, we expect similar solutions between the steps $t = 0.08s$ and $t = 0.16s$ and the steps $t = 0.12s$ and $t = 0.20s$. The reported pressure fields in Figure 1.16 for absorbing boundary conditions shows such behavior, in contrast to the outflow boundary condition.

We recall that, since this model is frequently used to simulate hemodynamic systems, the periodicity of the system is important because the blood vessels have usually a periodic behavior. In order to investigate further this point, we report in Figure 1.17 the pressure field along the cylinder axis for different, in-phase time steps. In order to guarantee the periodic conditions in the studied channel, we expect the same solutions at each of the reported time steps. In particular, it can be seen that the solutions with absorbing boundary conditions are the same at each studied step, and allows the exit of the pressure wave through the outlet of the channel. Instead, in the outflow case the spurious reflections don't allow to have the same profile in the in-phase steps.

Table 1.9: L^2 norm of the pressure over the entire domain at various time steps in the absorbing case for in-phase steps.

Time $[s]$	$\|p\|_{abs}$
0.26	252.72
0.34	253.87
0.42	254.82
0.50	254.90

In Table 1.9 the comparison between the L^2-norm of the pressure in the absorbing case for the four studied steps is reported. The global pressure field shows a small discrepancy between each analyzed in-phase step. In conclusion, the absorbing boundary conditions are necessary to have a correct numerical simulation of periodic physical systems. This approach can be applied also to simple Navier-Stokes simulations.

Optimal boundary control models

The interest of the scientific community towards optimization problems, together with the growing computational power available, brings to a wide literature on optimization processes. Such works can be based on many different approaches, such as linear feedback methods, adjoint or sensitivities-based optimal control, multi-objective optimization and many others. The interested reader can consult [66, 67, 68, 69]. Linear feedback methods are commonly used to analyze complex systems, e.g. turbine valves and heat exchangers in power plants, and are also used in electronic applications. Multi-objective optimization and sensitivities-based optimal control are other interesting research fields with various applications, ranging from engineering design to financial predictions of market shares.

In this chapter an adjoint based method is used. Such a method has been proven to be a good approach for the optimal control of complex computational fluid dynamics problems [70]. Moreover, these methods have a solid mathematical background and the existence of local optimal solutions can be proven for many interesting cases [71]. However, this method is only appropriate when the design variables are continuous. Since the core of the control problem is the minimization of an objective function, in some cases such a function may contain multiple minima. In this case the classical gradient

approach will generally converge to the nearest minima, without searching
for other minima elsewhere in the design space. If the objective function
is known to have multiple local minima, and possibly discontinuities, then
a stochastic search method may be more appropriate. The adjoint based
method presented in this work can then be used only to get improvements
from a reference state and not to find the global optimal solution to the prob-
lem, unless this is the only minimum of the functional. However, in many
practical situations an improvement on a reference state is what is needed
because too big changes on the design cannot be performed for physical or
practical reasons.

The optimal control problems can be categorized in *stationary* problems,
where the system variables do not depend on time, and in *unsteady* problems,
where the studied system is optimized during its evolution in time. Another
classification can be made between different approaches to control, such as
distributed, *boundary* and *shape* controls. The difference between them lies
in the way the control can act on the problem domain. In the distributed
control problems source terms in the interior part of the domain are used as
control parameters. This kind of control is not often suitable for practical
applications due to technical difficulties to control the internal points of a do-
main. However, from a numerical point of view is usually easy to implement.
On the other hand, in the boundary control one acts on the system through
its external surface. It can be considered as a more realistic approach to
optimization. However, the mathematical approach to such controls is more
challenging, as well as the numerical algorithm implementation. Lastly, in the
shape control the controlled parameter is the shape of the domain boundary.
Furthermore, identification of material properties such as Young modulus in
solids or fluid viscosity is an inverse problem that can be studied with the
adjoint based formalism, see, for example, [72, 11].

In the following sections of this chapter, the basic principles of the adjoint
optimal control theory are presented by deriving the optimality system for a
simple example problem. Then, the Koiter fluid-structure interaction model
introduced in the last chapter is applied to optimal control problems. For this
purpose, we follow a mathematical formulation in order to show the existence
of the solution of the coupled fluid-structure system presented in the last
chapter. At the same time, we show the existence of the optimal solution,
for a given set of boundary conditions and under certain hypotheses. By
using the Koiter model, it is possible to study the FSI steady optimal control

problem without taking into consideration the real extension of the solid domain and thus simplify both the theoretical treatment about the optimal solution existence, regularity and differentiability properties.

Note that while optimal control problems of classical fluid-structure interaction are studied in many works in literature, there are no works on the optimal control of the Koiter fluid-structure interaction. Thus, the mathematical models presented in this chapter constitutes a novelty in the field of the optimal control.

2.1 A simple adjoint optimal control problem

We now introduce a simple example of an adjoint Dirichlet boundary optimal control problem, useful to introduce and clarify the main aspects of the method. To set up the problem we first need to choose the goal to reach and how we intend to do so. We introduce a cost functional, a mathematical formulation that measures how far from the desired target the studied system is. This is usually expressed in terms of the state system variables. We
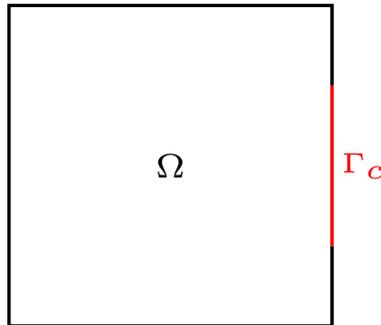
Figure 2.1: Domain Ω with control boundary Γ_c.

denote the control boundary as $\Gamma_c \subseteq \partial\Omega$ (see Figure 2.1). On Γ_c, Dirichlet optimal conditions are sought. Then, the optimal value of q on Γ_c to have a desired $u = u_d$ on the domain Ω_d, constrained by the simple equation $\Delta u = 0$ on Ω can be written as in the following problem.

Problem 2.1. *Find a state-control pair $(u, q) \in H^1(\Omega) \times Q$ which minimizes the cost functional*

$$\mathcal{J}(u, q) = \frac{1}{2}\|u - u_d\|_{L^2(\Omega_d)}^2 + \frac{\alpha}{2}G(q)^2, \tag{2.1}$$

under the constraints

$$(\nabla u, \nabla v) = 0 \qquad\qquad \forall v \in H_0^1(\Omega)\,, \qquad\qquad (2.2)$$

$$u = q \qquad\qquad on\ \Gamma_c\,, \qquad\qquad (2.3)$$

$$u = 0 \qquad\qquad on\ \partial\Omega \smallsetminus \Gamma_c\,. \qquad\qquad (2.4)$$

In this work, time-dependent optimal control problems are not considered. However, interested reader in such optimal control problems can see [67]. The cost functional (2.1) depends on the distance in norm between the solution u and the target field u_d, and on a regularization parameter $G(q)$. The aim of the method is to minimize $\mathcal{J}$ under the constraint of the equation of the state of the studied physical system (2.3). A typical example of regularization term is the L^2 norm $G(q) = \int_{\Gamma_c} q^2\, dx$.

2.1.1   The Lagrange multiplier method

The problem 2.1 may be reformulated as *unconstrained* optimization problems through the Lagrange multiplier method. We now introduce the definition of *local minimum*.

Definition 2.1. *A local minimum (u^*, q^*) for a functional $\mathcal{J}(u, q)$ is a point such that, for some $\varepsilon > 0$*

$$\mathcal{J}(u^*, q^*) \le \mathcal{J}(u, q) \qquad \forall (u, q) \in U \times Q\ and\ \|u^* - u\| < \varepsilon\,.$$

A Lagrange multiplier method is now applied in order to find the optimal solution of the presented problem. In order to do so we introduce the map $M : \boldsymbol{B}_1 \to \boldsymbol{B}_3$, with $\boldsymbol{B}_1 = U \times Q = H^1(\Omega) \times H^{1/2}(\Gamma_c)$ and $\boldsymbol{B}_3 = H^{-1}(\Omega) \times H^{-1/2}(\Gamma_c) \times H^{-1/2}(\Gamma_D)$, where $\Gamma_D = \Gamma \smallsetminus \Gamma_c$ We define M by $M(u, q) = (l_1, l_2)$ if and only if

$$(\nabla u, \nabla v) = (l_1, v) \qquad\qquad \forall v \in H_0^1(\Omega)\,,$$

$$\int_{\Gamma_D} u\phi\, dx = \int_{\Gamma_D} l_2\phi\, dx \qquad \forall \phi \in H^{1/2}(\Gamma_D)\,, \qquad\qquad (2.5)$$

Lemma 2.1. *Let the nonlinear mapping $M : \boldsymbol{B}_1 \to \boldsymbol{B}_3$ be defined by (2.5). We consider also the bounded linear operator $M'(u, q) : \boldsymbol{B}_1 \to \boldsymbol{B}_2$, with $\boldsymbol{B}_2 = H^{-1}(\Omega) \times H^{-1/2}(\Gamma_c) \times H^{-1/2}(\Gamma_D)$, and $M'(u, q) \cdot (\delta u, \delta q) = (\bar{l}_1, \bar{l}_2)$ for*

$(\delta u, \delta q) \in \boldsymbol{B}_1$ and $(\bar{l}_1, \bar{l}_2) \in \boldsymbol{B}_2$. Thus, the mapping M is strictly differentiable at a point $(u, q) \in \boldsymbol{B}_1$ and its strict derivative is given by M' if and only if

$$
\begin{aligned}
(\nabla \delta u, \nabla v) &= (\bar{l}_1, v) && \forall v \in H_0^1(\Omega), \\
\int_{\Gamma_D} \delta u \phi \, dx &= \int_{\Gamma_D} \bar{l}_2 \phi \, dx && \forall \phi \in H^{1/2}(\Gamma_D),
\end{aligned}
\tag{2.6}
$$

We consider now the operator $\mathcal{J}$ defined in (2.1) with the regularization term $G(q)$ defined by

$$
G(q) = \int_{\Gamma_c} q^2 \, dx,
$$

so we can define $\mathcal{J}$ as

$$
\mathcal{J} = \frac{1}{2} \int_{\Omega_d} (u - u_d)^2 \, d\boldsymbol{x} + \frac{\alpha}{2} \int_{\Gamma_c} q^2 \, dx
\tag{2.7}
$$

and its derivative as

$$
\mathcal{J}' = \int_{\Omega_d} \delta u (u - u_d) \, d\boldsymbol{x} + \alpha \int_{\Gamma_c} q \delta q \, dx.
\tag{2.8}
$$

Now, given a set $(u_1, q_1) \in U \times Q$, we can define a nonlinear mapping based on (2.5) and (2.7) as $Q : \boldsymbol{B}_1 \rightarrow \mathbb{R} \times \boldsymbol{B}_3$. For $a \in \mathbb{R}$, we have $Q(u, q) = (a, l_1, l_2, l_3)$ if and only if

$$
\begin{pmatrix} \mathcal{J}(u, q) - \mathcal{J}(u_1, q_1) \\ M(u, q) \end{pmatrix} = \begin{pmatrix} a \\ (l_1, l_2, l_3) \end{pmatrix}.
\tag{2.9}
$$

We can now introduce the following lemmas.

Lemma 2.2. *Let the nonlinear mapping $Q : \boldsymbol{B}_1 \rightarrow \mathbb{R} \times \boldsymbol{B}_3$ be defined by (2.9). The strict derivative of Q at a point (u, q) is given by Q' if and only if*

$$
\begin{pmatrix} \mathcal{J}'(u, q) \cdot (u, q) \\ M'(u, q) \cdot (\delta u, \delta q) \end{pmatrix} = \begin{pmatrix} \delta a \\ (\bar{l}_1, \bar{l}_2, \bar{l}_3,) \end{pmatrix}.
\tag{2.10}
$$

Lemma 2.3. *Let $(u, q) \in \boldsymbol{B}_1$ be a solution of the optimal control problem. We have that*

1. *the operator $M'(u, q)$ has closed range in $\boldsymbol{B}_2$,*

2. *the operator $Q'(u, q)$ has closed range in $\mathbb{R} \times \boldsymbol{B}_2$,*

3. *the operator $Q'(u,q)$ is not onto in $\mathbb{R} \times \boldsymbol{B}_2$.*

   Interested reader can find the proof of the Lemma 2.2 applied to a more complex mathematical problem in [73] and reference therein. Since the operator $Q'(u,q)$ is not onto, the first-order necessary condition is straightforward, see, e.g., [74]. Then, we can introduce the following theorem.

Theorem 2.1. *There exist a nonzero Lagrange multiplier $(\Lambda, u_a, q_a) \in \mathbb{R} \times \boldsymbol{B}_2^*$ satisfying*

$$\Lambda \mathcal{J}'(u,q) \cdot (\delta u, \delta q) + \Big\langle (u_a, q_a), M'(u,q) \cdot (\delta u, \delta q) \Big\rangle = 0 \quad \forall (\delta u, \delta q) \in \boldsymbol{B}_1 \quad (2.11)$$

where $< \cdot, \cdot >$ denotes the duality pairing between $\boldsymbol{B}_2$ and its dual space $\boldsymbol{B}_2^*$.

2.1.2   The optimality system

We can now derive the optimality system for the presented operator. In particular, starting from the equation (2.11) we can derive the following theorem.

Theorem 2.2. *Let $(u,q) \in \boldsymbol{B}_1$ denote a solution of the optimal control problem. Therefore, the control q is solution of the simple scalar equation*

$$q = -\frac{\nabla q_a \cdot \boldsymbol{n}}{\alpha}, \qquad on \ \Gamma_c, \tag{2.12}$$

while the variable u_a is the solution of the equation

$$\int_{\Omega_d} (u - u_d) v_a \, d\boldsymbol{x} - \int_{\Gamma_c} \nabla u_a \cdot \boldsymbol{n} v_a \, dx + (\nabla u_a, \nabla v_a) = 0, \tag{2.13}$$

with the boundary equation

$$\int_{\Gamma_D} u_a \phi_a = 0. \tag{2.14}$$

Proof. We are free to consider $\Lambda = 1$. Therefore, we can rewrite the equation (2.11) considering (2.8) and (2.6) as

$$\left(\int_{\Omega_d} (u - u_d) \delta u \, d\boldsymbol{x} + \alpha \int_{\Gamma_c} q \delta q \, dx \right) + (\nabla \delta u, \nabla u_a) + \int_{\Gamma_D} \delta u q_a \, dx = 0,$$

by integrating the term $(\nabla \delta u, \nabla u_a)$ by parts we obtain

$$\begin{aligned} &\left(\int_{\Omega_d} (u - u_d) \delta u \, d\boldsymbol{x} + \alpha \int_{\Gamma_c} q \delta q \, dx \right) + \int_{\Gamma_c} \delta q \nabla q_a \cdot \boldsymbol{n} \, dx \\ &\quad - (\delta u, \Delta u_a) + \int_{\Gamma_D} \delta u q_a \, dx = 0, \end{aligned} \tag{2.15}$$

Then, by imposing $\delta u = 0$ we can obtain

$$\int_{\Gamma_c} (\alpha q + \nabla q_a \cdot \boldsymbol{n}) \delta q \, dx \,.$$

Since δq is arbitrary, we can find the *optimality condition* (2.12). Moreover, by imposing $\delta q = 0$ to (2.15) we obtain

$$\int_{\Omega_d} (u - u_d) \delta u \, d\boldsymbol{x} - (\delta u, \Delta u_a) = 0$$

together with the boundary condition (2.14). Since δu is arbitrary, we can use it as a test function v_a. Thus, we can now integrate again by part in order to obtain the equation (2.13). Note that the variation δu and δq are used as test functions for the obtained equations. $\qquad\square$

In the following, we call u_a and q_a as *adjoint variables*. Note that the adjoint variable u_a in (2.13) satisfies an equation similar to the starting one (2.2). Moreover the boundary conditions of the adjoint system are of the same type as in the state problem but they should be assumed homogeneous. This is a typical feature of adjoint problems: when a Dirichlet boundary condition is imposed in the state system, a homogeneous Dirichlet boundary condition must be imposed in the adjoint system.

The optimality system is composed by the *state equation*, the *adjoint equation* and the *optimality condition*. The state equation is the initial constraint (2.1) together with the boundary conditions (2.2). The adjoint system is composed by the equation (2.13) together with the boundary condition (2.14). The optimality condition is reported in (2.12), as reported above.

All the presented equations form the so-called *optimality system* and allow finding the stationary point of the Lagrangian functional, and therefore to find the optimal solution of the presented problem.

2.1.3 Numerical solution of the optimality system

To solve the optimality system introduced in the last section, one can use a *one shot* method. In that case, the system is solved in a fully coupled fashion. Using this approach, in the example introduced in the last sections the system is solved in a fast and reliable way without any optimization algorithm. However, for complex mathematical problems the optimality system is made of many nonlinear and strongly coupled equations. Moreover the solution becomes too expensive when a fine domain discretization is used. Under these

hypotheses the one-shot method becomes unattractive. In order to overcome these challenges, a *segregated* approach is generally a better alternative. The segregated method provides that all the equations of the optimality system are solved independently with appropriate algorithms, known as optimization algorithms, that reduce the computational cost of the whole simulation. Both the presented methods require a starting point x_0. Starting from it, a sequence of solutions $\{x_k\}_{k=0}^{\infty}$ is found and the algorithm ends when either a certain convergence criteria are satisfied.

In literature some classes of algorithms can be found, that differ in the way the updated solutions are found from the previous ones. We consider now the *trust region* and the *line search* methods. Both of them require a search direction and a step size, which evaluate the distance from the current solution. The trust region method provides that every new iterate is searched in a region around the current solution. The maximum distance between two consecutive iterates is fixed by a step size and the algorithm aims to find the optimal direction. The line search method provides that the step direction is found first and then the step size is chosen in order to minimize the functional along that direction.

In this work we use a gradient based line search approach with a simple Armijo backtracking strategy [75]. The search direction is obtained by solving the adjoint equations and the optimality condition that gives the objective functional gradient direction. A generic iteration of a line search method for the minimization of a functional $\mathcal{J}$ is given by

$$x_{k+1} = x_k + \alpha_k p_k \,,$$

where k is the current iteration, p_k is the search direction and α_k is the step length. The search direction has to be chosen in the descent direction. For this reason it usually can be written as

$$p_k = -B_k^{-1} \nabla \mathcal{J}_k \,,$$

where B_k is a symmetric, non-singular, positive definite matrix. The choice of B_k defines the used method. For example, when $B_k = I$ (the identity matrix) the search direction is given by the functional gradient and the method is known as *steepest descent*. In Newton methods B_k is the Hessian matrix of the functional $\Delta \mathcal{J}$, while in quasi-Newton methods B_k is an approximation of the Hessian at every iteration with a low-rank formula, such as BFGS [76].

The step length parameter α_k can be chosen in different ways. An *exact line search* aims to find the value of α_k that minimizes the functional $\mathcal{J}_k$

Algorithm 2 Backtracking Line search

---
    1. Set $\alpha_k = \alpha_0 > 0$, $\rho \in (0, 1)$
  while $\mathcal{J}_k(x_k + \alpha_k p_k) > \mathcal{J}_k(x_k)$ **do**
      2. Set $\alpha_k = \rho\,\alpha_k$
    if $\alpha_k < \varepsilon$ **then**
        Line search not successful   ▷ *End of the algorithm*
    end if
  end while

---

along the search direction. This method significantly reduces the value of the functional, however is a computationally expensive algorithm. For this reason more feasible strategies perform only an *inexact line search*, reducing $\mathcal{J}_k$ at minimal cost. The *backtracking* line search strategy presented in Algorithm 2 is a commonly used inexact method. The step length α_k is first initialized to a positive value α_0, whose value depends on the choice of the algorithm. Then the step length is reduced by a contraction factor ρ until a lower value of the functional for the new iterate is found. When $\alpha_k < \varepsilon$ the process ends.

A final brief remark on the sensitivity of the functional is now needed. In the numerical approach to optimal control problems, it is very important to handle control or design parameters such that small changes in their values imply large changes in the solution. This means that the solution should be very sensitive to small changes in the data. The insensitivities of the cost functional can be used to induce changes in the choice of design parameters by replacing the useless parameters with others that have a greater effect on the cost functional, or change the cost functional itself so that it becomes more sensitive to the design parameters.

All the numerical computations of optimal control simulations introduced in this chapter will be done with the numerical approach presented in this section. In the final chapter, a different approach based on a monolithic system will be used.

2.2 Optimal boundary control of steady Koiter FSI model

In this section, an optimal boundary control for the steady Koiter FSI model presented in the last chapter is introduced. We follow an approach similar to the technique introduced to obtain the optimality system for the problem

2.1. In particular, we consider as a controlled variable the pressure on the domain Γ_c. The presented formulation also accounts the possibility to control the velocity field on Γ_c, defining two different classes of problem, the pressure boundary control, and the velocity boundary control. The goal of the presented control is to obtain a desired displacement field η_d over a domain $\Omega_d \subseteq \Omega$. In order to do so, we follow a mathematical formulation similar to the approach used in [73, 71]. However, the moving domain and the mesh motion complicate the mathematical and numerical approach to the control problem.

In order to maintain a consistent mathematical formulation of the optimal control problem, we consider a Koiter shell model slightly different from the model presented in the last chapter, following the model introduced in [77]. Therefore, a brief introduction of the modified model is needed, together with some basic notation useful in the rest of this section.

2.2.1   An introduction to the FSI control problem

In this work, the membrane model based on the Koiter shell equations is used to reduce the space dimension of the solid structure [4]. Let $\Omega \subset \mathbb{R}^2$ be the domain where Koiter fluid-structure interaction equations are solved, with boundary Γ. As noted in the previous chapter, using this approach the solid region of a full classical FSI model collapses into the fluid-solid boundary Γ_s. Following the notation in Figure 2.2 on the left, the reference domain is bounded by Γ_2 (inlet), Γ_4 (outlet), Γ_3 (no-slip wall) and Γ_0 (fluid-solid reference boundary). We also call A and B the extreme points of Γ_0. In this section, $\hat{\boldsymbol{n}}$ will denote the unit vector normal to the boundary Γ.

The optimal control theory may face difficulties when dealing with moving domains. For this purpose, since we often need in this section to compute and differentiate quantities on moving boundaries, an extension of the functions on an extended fixed domain is required in order to keep the mathematical formulation consistent. In the rest of this section, we use the $(\widehat{\cdot})$ notation for quantities extended on a fixed domain. For example, given a function u defined over the moving domain Ω, we denote with $\widehat{u}$ the extension on the extended domain $\widehat{\Omega}$ of u. Thus, we introduce the Calderon extension theorem in order to build properly the extension functions.

Theorem 2.3. *Let $\Omega \subset \mathbb{R}^2$ be a uniform Lipschitz domain, and E a linear*
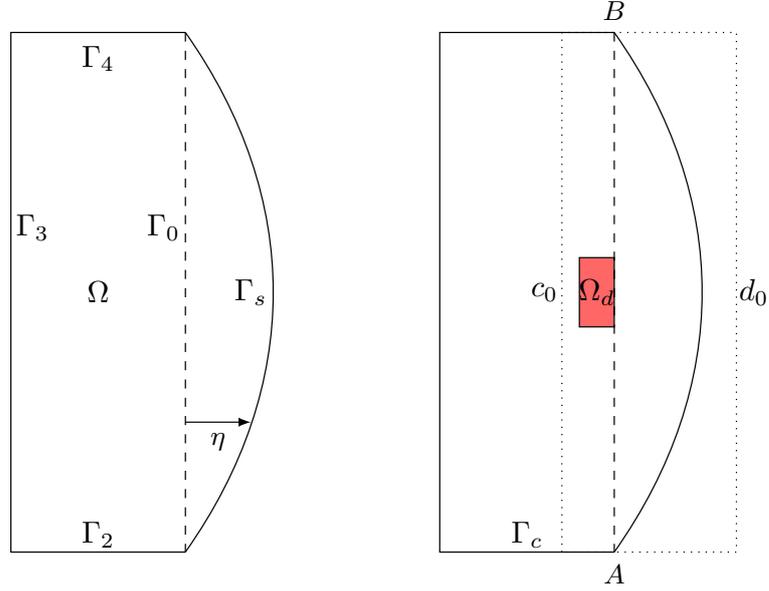
Figure 2.2: On the left, the FSI Koiter domain with the name of all boundaries. On the right, some parameters useful for the optimal control are presented.

continuous extension operator

$$E \ : \ \boldsymbol{H}^s(\Omega) \to \boldsymbol{H}^s(\mathbb{R}^2) \,.$$

Then, $\forall \boldsymbol{u} \in \boldsymbol{H}^s(\Omega) , \exists E$ such that $\|E\boldsymbol{u}\|_m \leq K\|\boldsymbol{u}\|_m$, where the constant $K \in \mathbb{R}^+$ depends only on the cone embedded in Ω.

Now some basic notations on the moving domain are introduced. The solid region, which collapses into the domain boundary Γ_1, can be mapped over an interval (a, b) denoted by I. We introduce r as a function defined over Ω representing the projection of η on from I to Γ_1 and extended inside of the domain, as

$$a(r, \phi) = 0 \qquad \forall \phi \in H_0^1(\Omega_\eta) \,.$$

In this case, with reference to Figure 2.2 we can state that the coordinates of the solid domain can be written as

$$\begin{aligned}
\Gamma_1 = \{ \boldsymbol{x} \in \mathbb{R}^2 \,|\, \boldsymbol{x} = (x(t), y(t)) \text{ with } t \in I = (a, b), \\
x = \eta(t), y = y(A) + y(B)(t - a)/(b - a) \} \,.
\end{aligned} \tag{2.16}$$

Under the condition (2.16), in the following the domain Ω and Γ_1 can be denoted by $\Omega(r)$ and $\Gamma_1(r)$. However, for the sake of simplicity, in the following

the dependence on r is omitted. In order to enable to explicitly derive a first order necessary condition of the optimal problem, the function $r(t)$ shall be regular enough. In order to do so, we shall require that the fluid-solid boundary Γ_1 has to be $C^{1,1}$ [78]. Moreover, to have a regular flow the domain must have convex corners. Under these hypotheses (Γ_1 has to be $C^{1,1}$ and Ω with convex corners), we can state that the membrane Γ_1 is connected smoothly to the rest of the boundary.

The reference configuration of the fluid-solid boundary Γ_0 can be defined as the set of points $\boldsymbol{x}_0(t) = (\eta_0(t), t) \in \mathbb{R}^2$ with $t \in I = [0, L]$. As can be seen in Figure 2.2, we consider a line parallel to the y-axis of length L as the reference undeformed configuration. The tangent and the normal unit to Γ_0 will be denoted by $\boldsymbol{t}_0$ and $\boldsymbol{n}_0$, respectively. The displacement $\eta(t) : I \to \mathbb{R}$ defines the deformed boundary $\Gamma_1(t)$ as the set of points $\boldsymbol{x}_1(t) = \boldsymbol{x}_0(t) + \boldsymbol{n}_0(t)\eta(t) \in \mathbb{R}^2$ depending on the parameter $t \in I$ defined in (2.16). When the reference configuration $\boldsymbol{x}_0$ is a line parallel to the y-axis, we have $\boldsymbol{x}_0(t) = (\eta_0(0), t)$, $\boldsymbol{n}_0 = \widehat{i}_x$, $\boldsymbol{t}_0 = \widehat{i}_y$ and $\boldsymbol{x}_1(t) = (\eta(t), t) = (x_0(0) + \eta(t), t)$ with $\boldsymbol{t}_0 = (t_{0x}, t_{0y}) = (\eta'/(\sqrt{1 + \eta'^2}), 1)$, $\boldsymbol{n}_0 = (t_{0y}, -t_{0x})$ and $t \in [0, L]$.

The physical model

We consider an incompressible viscous fluid flowing through the domain Ω, as shown in Figure 2.2. The velocity $\boldsymbol{u}$ and the pressure p satisfy the stationary incompressible Navier-Stokes system

$$-\nu\nabla \cdot \boldsymbol{D} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} + \nabla p = \boldsymbol{f} \qquad \text{on} \quad \Omega\,, \qquad (2.17)$$

$$\nabla \cdot \boldsymbol{u} = 0 \qquad \text{on} \quad \Omega\,, \qquad (2.18)$$

where ν the constant fluid viscosity, $\boldsymbol{D}(\boldsymbol{u}) = (\nabla \boldsymbol{u}^T + \nabla \boldsymbol{u})/2$ is the deformation tensor and $\boldsymbol{f}$ is a body force. The velocity vanishes along the walls where no-slip boundary conditions are imposed. In particular, a no-slip condition is imposed on the fixed wall Γ_3 and on the moving boundary Γ_1.

The solid behavior is modeled with the Koiter membrane model studied in section 1.4. As mentioned above, the model relies on some assumptions: small deformations, negligible bending term, linear constitutive stress-strain relation, and only normal displacement. Under these hypotheses, in the current rectangular case (as reported in Figure 2.2 on the left) the displacement field $\boldsymbol{\eta}$ can be written as $\boldsymbol{\eta} = (\eta, 0)$. Following the procedure in [5] and, in particular, in [77], the stationary Koiter model on the reference configuration

Γ_0 can be written as

$$\beta\eta - \gamma\Delta\eta - \gamma_1\Delta^2\eta = f_s \qquad\qquad \text{on } \Gamma_0\,, \qquad\qquad (2.19)$$

where β, γ and $\gamma_1 \in \mathbb{R}$ can be considered constants, under some simplifying assumptions valid for simple geometries (e.g. the considered rectangle). In particular, β depends on the Young modulus of elasticity E, the Poisson ratio and the geometry of the solid; γ depends on the prestress state of the solid material, e.g. for simplified geometry it depends on the solid viscosity. For $\gamma = \gamma_1 = 0$ we recover the pure elastic case. Note that, in contrast to some works in literature, a fourth-order term has been added to the equation, accordingly to [79, 80]. Now, in order to couple the fluid-solid with the same test functions we introduce a new variable κ such that

$$\gamma\Delta\eta + \gamma_1\Delta\kappa - \beta\eta + f_s = 0 \qquad\qquad \text{on } \Gamma_0\,, \qquad\qquad (2.20)$$
$$\kappa = \Delta\eta \qquad\qquad \text{on } \Gamma_0\,. \qquad\qquad (2.21)$$

Therefore, the numerical physical system is built in such a way that the displacement field η (defined on Γ_0) must satisfy the Navier-Stokes system (2.17)-(2.18) on Ω, and the Koiter system (2.20)-(2.21) on Γ_0.

The optimal control problem

We can now introduce the optimal control problem studied in this section.

Problem 2.2. *Find a state-control set $(\tau_c, \boldsymbol{g}_c, \eta) \in L^2(\Gamma_c) \times \boldsymbol{H}^1(\Gamma_c) \times H^2(\Gamma_0)$ which minimizes the cost functional*

$$\mathcal{J}(\tau_c, \boldsymbol{g}_c, \eta) = \frac{1}{2}\int_{\Omega_d} |\eta - \eta_d|^2 dx + \frac{\lambda}{2}\|\tau_c\|_{L^2(\Gamma_c)} + \frac{\lambda_1}{2}\|\boldsymbol{g}_c\|_{H^1(\Gamma_c)}\,, \qquad (2.22)$$

where $\lambda, \lambda_1 \in \mathbb{R}^+$ are constant, and under the constraints

$$-\nu\nabla\cdot\boldsymbol{D} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p = \boldsymbol{f} \qquad\qquad on \quad \Omega\,,$$
$$\nabla\cdot\boldsymbol{u} = 0 \qquad\qquad on \quad \Omega\,,$$
$$\gamma\Delta\eta + \gamma_1\Delta\kappa - \beta\eta + f_s = 0 \qquad\qquad on \quad \Gamma_0\,,$$
$$\kappa = \Delta\eta \qquad\qquad on \quad \Gamma_0\,,$$

and a consistent set of boundary conditions for all the involved variables (u, p, η and κ).

In other words, the optimal control problem consists in finding the optimal pressure field on Γ_c (pressure boundary control, $\lambda \neq 0$ and $\lambda_1 = 0$) or the optimal velocity field on Γ_c (velocity boundary control, $\lambda_1 \neq 0$ and $\lambda = 0$) in order to have a desired displacement η_d on the domain Ω_d. Since the presented problem is different from the standard optimal control problems available in literature, some additional requirements have to be imposed in order to show the existence of the optimal solution of the numerical problem and the consistency of the Lagrange multiplier method used to solve the optimal control problem.

2.2.2   Weak formulation of the constraints

Now we introduce the weak formulation of the constraints of the problem 2.2 together with a set of consistent boundary conditions. The multi-scale fluid-structure problem can now be introduced. We introduce $\boldsymbol{g} \in H^{1/2}(\Gamma)$ as the boundary velocity satisfying the compatibility condition $\int_\Gamma \boldsymbol{g} \cdot \boldsymbol{n} ds = 0$, that vanishes on the fluid-solid boundary where no-slip boundary conditions are considered. The state $(\boldsymbol{u}, p, \eta, \kappa, r) \in \boldsymbol{H}^1(\Omega) \times L_0^2(\Omega) \times H_0^2(I) \times H_0^1(I) \times H^2(I)$ is constrained to satisfy the fluid system (2.17)-(2.18) and the Koiter shell system (2.20)-(2.21), together with the appropriate boundary conditions. By using appropriate test functions $\boldsymbol{\varphi}$ the split weak form of fluid-structure problem can be obtained. Given the normal displacement and the boundary parametrization (η, κ, r), the velocity and pressure fields $(\boldsymbol{u}, p)$ must satisfy the following weak system

$$
\begin{aligned}
\nu a(\boldsymbol{u}, \boldsymbol{\varphi}) + c(\boldsymbol{u}; \boldsymbol{u}, \boldsymbol{\varphi}) + b(\boldsymbol{\varphi}, p) &= (\boldsymbol{f}, \boldsymbol{\varphi}) && \forall \boldsymbol{\varphi} \in \boldsymbol{H}_\Gamma^1(\Omega(r)), \\
b(\boldsymbol{u}, \psi) &= 0 && \forall \psi \in L_0^2(\Omega(r)), \\
(\boldsymbol{u}, \boldsymbol{s})_\Gamma &= (\boldsymbol{g}, \boldsymbol{s})_\Gamma && \forall \boldsymbol{s} \in \boldsymbol{H}^{-1/2}(\Gamma(r)).
\end{aligned}
\tag{2.23}
$$

Moreover, the solid system reads

$$
\begin{aligned}
(f_s, \psi)_I + (\beta\eta, \psi)_I + (\gamma\nabla\eta, \nabla\psi)_I + (\gamma_1\nabla\kappa, \nabla\psi)_I &= 0 & \forall \psi \in H_0^1(I), \\
(\kappa, \psi_1)_I &= (\gamma_1\nabla\eta, \nabla\psi_1)_I & \forall \psi_1 \in H_0^1(I),
\end{aligned}
\tag{2.24}
$$

where we assume Dirichlet boundary conditions on both η and κ at the extreme points of the interval I. In the following, we refer to $\Omega = \Omega(r) = \Omega_r$ by using these notations.

The coupling between the system for the fluid (2.23) and the system for the solid (2.24) is obtained through the normal stress tensor $\boldsymbol{\tau}_n = -p\boldsymbol{n} +$

$\nu\big(\nabla\boldsymbol{u}+(\nabla\boldsymbol{u})^T\big)\cdot\boldsymbol{n}$. In fact, we may impose $(\boldsymbol{\tau}_n,\boldsymbol{n}_0\psi)_{\Gamma_1}=(\boldsymbol{\tau}_n,\boldsymbol{n}_0\,|J_1^{-1}|\,\psi)_I=(f_s,\psi)_I$, where J_1 is the Jacobean of the transformation from the reference configuration I to the deformed configuration Γ_1. For numerical purpose, it is more convenient to write all the equations over the deformed configuration. Then, by introducing the functional space

$$\boldsymbol{H}_{t_0}^1(\Omega(r))=\{\boldsymbol{f}\in\boldsymbol{H}^1(\Omega(r))|\,\boldsymbol{f}-\boldsymbol{n}_0(\boldsymbol{n}_0\cdot\boldsymbol{f})=0\text{ on }\Gamma_1\}\,,\qquad(2.25)$$

we can obtain the following coupled system

$$\begin{aligned}
\nu a(\boldsymbol{u},\boldsymbol{\varphi})+c(\boldsymbol{u};\boldsymbol{u},\boldsymbol{\varphi})&+b(\boldsymbol{\varphi},p)+(\boldsymbol{\tau}_n,\boldsymbol{\varphi})_{\Gamma_N}\\
-\,(\beta\eta,\varphi_{n_0}J_1)_{\Gamma_1}&-(\gamma J_1\nabla_{t_0}\eta,\nabla_{t_0}\varphi_{n_0})_{\Gamma_1}\\
-\,(\gamma_1 J_1\nabla_{t_0}\kappa,\nabla_{t_0}\varphi_{n_0})_{\Gamma_1}&=(\boldsymbol{f},\boldsymbol{\varphi})&&\forall\boldsymbol{\varphi}\in\boldsymbol{H}_{t_0}^1(\Omega)\,,\\
b(\boldsymbol{u},\psi)&=0&&\forall\psi\in L^2(\Omega)\,,\\
(\kappa,\psi_1)&=(\nabla_{t_0}\eta,\nabla_{t_0}\psi_1)&&\forall\psi_1\in H_0^1(I)\,,\\
(\boldsymbol{u},\boldsymbol{s})_\Gamma&=(\boldsymbol{n}_0 g_n,\boldsymbol{s})_\Gamma&&\forall\boldsymbol{s}\in\boldsymbol{H}^{-1/2}(\Gamma)\,,\\
a(r,\phi)&=0&&\forall\phi\in H_0^1(\Omega)\\
(r,\phi_1)_\Gamma&=(J_1\eta,\phi_1)_{\Gamma_s}-(\boldsymbol{u}\cdot\boldsymbol{n},\phi_1)_{\Gamma_s}&&\forall\phi_1\in L^2(\Gamma)\,,
\end{aligned}\qquad(2.26)$$

where Γ_N represents the boundary where Neumann conditions are imposed.

Note that the presented weak form (2.26) balances the fluid normal stress along the $\boldsymbol{n}_0$ direction with the solid force

$$\big(-p\hat{\boldsymbol{n}}+\nu(\nabla\boldsymbol{u}+(\nabla\boldsymbol{u})^T)\cdot\hat{\boldsymbol{n}}\big)_{\Gamma_1}=\big(\beta\eta-\gamma\nabla^2\eta-\gamma_1\nabla^4\eta\big)$$

The tangential solid force (i.e. in the $\boldsymbol{t}_0$ direction) is balanced by the no-slip boundary condition that does not allow tangential displacements.

In the following, we consider the monolithic formulation (2.26) and the splitted formulation (2.23)-(2.24) as equivalent. Considering the split formulation, we can refer to the standard results in [43] in order to show the existence and the uniqueness of the solution. In particular, for the fluid equation (2.23), we can introduce the following Theorem.

Theorem 2.4. *Let Ω be an open, bounded set of $\mathbb{R}^2$ with Lipschitz-continuous boundary Γ. Let $\boldsymbol{f}\in\boldsymbol{H}^{-1}(\Omega)$ and $\boldsymbol{g}\in\boldsymbol{H}^{1/2}(\Gamma)$. Then, there exists at least one solution $(\boldsymbol{u},p)\in\boldsymbol{H}^1(\Omega)\times L_0^2(\Omega)$ of the system (2.23). The set of $\boldsymbol{u}$ is closed in $\boldsymbol{H}^1(\Omega)$ and is compact in $\boldsymbol{L}^2(\Omega)$.*

Moreover, for some positive $\nu_0=\nu_0(\Omega(r),\boldsymbol{f},\boldsymbol{g})$ depending on the given data, if $\nu>\nu_0$ then the set of solutions of (2.23) consists of a single element.

With reference to the same cited work, we can also obtain the corresponding Theorem for the existence of a smoother solution of (2.23). To introduce such a theorem, we rely on the validity of the inf-sup condition on the operator $B(\boldsymbol{u}, (q, d)) = b(\boldsymbol{u}, (q)) + (\boldsymbol{u}|_\Gamma, d)$ and the norm $\|\boldsymbol{g}\|_{1,\Gamma}$ (see [43, 9, 81]). Then, we consider the constants c_a and c_b that follows the coercivity properties (see [44])

$$a(\boldsymbol{u}, \boldsymbol{u}) \geq c_a \|\boldsymbol{u}\|_1^2 \qquad \boldsymbol{u} \in \boldsymbol{H}_{\Gamma_\alpha}^1(\Omega)\,,$$

$$\inf_{p \in L_0^2(\Omega)} \sup_{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)} \frac{b(\boldsymbol{v}, p)}{\|\boldsymbol{v}\|_1 \|p\|} \geq c_b\,.$$

Theorem 2.5. *Let $r \in H^2(I)$ and $\boldsymbol{\tau}_n \in \boldsymbol{H}^{-1/2}(\Omega)$. Thus, there exists at least one solution $(\boldsymbol{u}, p) \in \boldsymbol{H}^{3/2}(\Omega) \times H^{1/2}(\Omega) \cap L_0^2(\Omega)$ of the system (2.23).*

Moreover, for some positive constant $C_2 = C_2(\nu, c_b, k_a, k_c)$, which does not depend on the domain Ω and its boundary Γ, the solution satisfies the following estimates

$$\|p\|_{0,\Omega} + \|\boldsymbol{\tau}_s\|_{0,\Gamma} \leq C_2 \Big(\|\boldsymbol{f}\|_{0,\Omega} + \|\boldsymbol{g}\|_{1,\Gamma} + \|\boldsymbol{g}\|_{1,\Gamma}^2 \Big)\,. \qquad (2.27)$$

Now we consider the shell system (2.24): since the system is a fourth order mono-dimensional differential equation, the regularity comes from standard theory (see [43]). We also remark that, from the embedding Sobolev theorem, $\eta \in H^2(I) \subset C^0(I)$. Then, we can introduce the following Theorem.

Theorem 2.6. *Let $f_s \in L^2(I)$. There exists a solution $(\eta, \kappa) \in H^1(\Omega) \cap H_0^1(I) \times H^1(I)$ of the system (2.24).*

Furthermore, there exists a positive constant $C_1 = C_1(\beta, \gamma_1, \gamma)$ such that $\|\eta\|_3 \leq C_1 \|f_s\|$. If there exists a positive constant k_s such that $\|f_s\| \leq k_s$, then there is a positive constant C_2 such that $|\eta(t)| \leq C_2$, $\forall\, t \in I$.

We consider now the monolithic system (2.26). Such a system is equivalent to the presented split system. For the boundary velocity we set $\boldsymbol{u}|_\Gamma = \boldsymbol{t} g_t + \boldsymbol{n} g_n$. We assume that $g_t = 0$ and the inflow velocity $g_n \in H^1(\Gamma)$ is defined directly or through appropriate Neumann conditions. We now need to bound the force $\boldsymbol{f}$ independently of Γ_1. Let $\widetilde{\Omega}$ be an open bounded extension of the domain Ω. In this way $\|f_n\|_{\widetilde{\Omega}_f}$ bounds $\|f_n\|_\Omega$ for all Ω. Then, by using the Theorems 2.5 and 2.6 we can assert that the deformations are limited. Thus, all possible domains can be included inside an extended domain $\widehat{\Omega}$.

Now we reformulate the Navier-Stokes problem over the extended domain $\widehat{\Omega}$. We consider the variables $(\boldsymbol{u}, \boldsymbol{g}, p, \tau_n)$ of the system (2.23) for all possible

values (η, κ, r). Under these hypotheses, we can study the problem with fixed domain theorems. We remark that the $(\widehat{\cdot})$ notation indicates the quantities extended over $\widehat{\Omega}$. The construction of the extended function is obtained by Calderon's extension theorem (see Theorem 2.3).

Since the movement of the boundary Γ_1 is defined by the displacement η, which is bounded above and below as defined above, then we can consider a fixed domain $\widehat{\Omega}$ with boundaries $\widehat{\Gamma}$ and $\widehat{\Gamma} \cap \Gamma = \Gamma \setminus \Gamma_1$ that contains Ω. We consider $\tau_n \in \mathcal{T}$. There exists $\boldsymbol{g} \in \boldsymbol{H}^1(\Gamma)$ such that $(\boldsymbol{u}, p, \tau_n) \in \boldsymbol{H}^{3/2}(\Omega) \times H^{1/2}(\Omega) \times \mathcal{T}$ satisfies (2.23). Now we consider the following procedure to construct the extension $(\widehat{\boldsymbol{u}}, \widehat{p}, \widehat{\boldsymbol{f}})$ in $\boldsymbol{H}^1(\widehat{\Omega}) \times L^2(\widehat{\Omega}) \times \boldsymbol{L}^2(\widehat{\Omega})$ from the solution $(\boldsymbol{u}, p, \boldsymbol{f})$ in $\boldsymbol{H}^1(\Omega) \times L^2(\Omega) \times \boldsymbol{H}^{-1}(\Omega)$. We set $\widehat{\boldsymbol{u}} = \boldsymbol{u}$ and $\widehat{p} = p$ over Ω and zero over $\widehat{\Omega} \setminus \Omega$. By using this extension $\widehat{\boldsymbol{f}} = \boldsymbol{f}$ over Ω and $\widehat{\boldsymbol{f}} = 0$ over $\widehat{\Omega} \setminus \Omega$. The test functions can be naturally extended. With this setting, $(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{g}}, \widehat{p}, \tau_n)$ satisfies the Navier-Stokes system with velocity field and test functions that vanish on $\partial \widehat{\Omega}$, namely satisfies the following system.

$$
\begin{aligned}
\nu a(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + c(\widehat{\boldsymbol{u}}; \widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + b(\widehat{\boldsymbol{\varphi}}, \widehat{p}) &= (\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{\varphi}})_{\widehat{\Omega}} && \forall \widehat{\boldsymbol{\varphi}} \in \boldsymbol{H}^1_{\partial\widehat{\Omega}}(\widehat{\Omega}), \\
b(\widehat{\boldsymbol{u}}, \widehat{\psi}) &= 0 && \forall \psi \in L^2_0(\widehat{\Omega}), \qquad (2.28) \\
(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{s}})_{\partial\widehat{\Omega}} &= (\widehat{\boldsymbol{g}}, \boldsymbol{s})_{\partial\widehat{\Omega}} && \forall \widehat{\boldsymbol{s}} \in \boldsymbol{H}^{-1/2}(\partial\widehat{\Omega}).
\end{aligned}
$$

and the normal boundary stress can be computed as

$$
\begin{aligned}
(\boldsymbol{n}_0 \tau_n, \widehat{\boldsymbol{\varphi}})_{\partial\widehat{\Omega}} &= \nu a(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + c(\widehat{\boldsymbol{u}}; \widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + b(\widehat{\boldsymbol{\varphi}}, \widehat{p}) - (\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{\varphi}})_{\widehat{\Omega}} && \forall \widehat{\boldsymbol{\varphi}} \in \boldsymbol{H}^1_{\boldsymbol{t}_0}(\widehat{\Omega}), \\
(\boldsymbol{n}_0 \tau_n, \widehat{\boldsymbol{\varphi}})_{\Gamma_1} &= \nu a(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + c(\widehat{\boldsymbol{u}}; \widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + b(\widehat{\boldsymbol{\varphi}}, \widehat{p}) - (\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{\varphi}})_{\Omega} && \forall \widehat{\boldsymbol{\varphi}} \in \boldsymbol{H}^1_{\boldsymbol{t}_0}(\Omega), \\
&&& \qquad\qquad (2.29)
\end{aligned}
$$

or, when $\tau_n \in L^2(\partial\widehat{\Omega})$, in the proper explicit form, as

$$
\tau_n = -p\boldsymbol{n}_0 + \nabla\widehat{\boldsymbol{u}} \cdot \boldsymbol{n}_0.
$$

on $\partial\widehat{\Omega}$ or Γ_1. The stress τ_n on $\partial\widehat{\Omega} \cap (\Gamma - \Gamma_1)$ can be used for appropriate Neumann conditions and the stress τ_n on Γ_1 as driven force over the solid region. We remark that only $\tau_n \in \mathcal{T}$ results in boundary Neumann condition with pressure-velocity solutions in standard spaces. With the introduced basics, we can now state the existence of the solution of the fluid-structure coupled system.

Theorem 2.7. *Let $\Omega_0 \subset \mathbb{R}^2$ be the fluid reference open bounded domain with piece-wise $C^{1,1}$ boundary Γ_0. Also let $\boldsymbol{f} \in \boldsymbol{L}^2(\widetilde{\Omega}_f)$ and $g_n \in H^1(\Gamma)$ with $\widetilde{\Omega}_f \subset \Omega_0$ and $\boldsymbol{g}$ defined as above. Then there exists a domain Ω_η and*

a $(\boldsymbol{u}, p, \eta, \kappa) \in \boldsymbol{H}^{3/2}(\Omega_\eta) \times H^{1/2}(\Omega_\eta) \times H_0^1(I) \times H^1(I)$ *solution of the* (2.23) *and* (2.24).

Proof. In order to prove the presented Theorem, we consider the Schauder's fixed point theorem (2.23)-(2.24). First, note that for any (η, κ, r) in their consistent spaces, the terms $\boldsymbol{g} \in \boldsymbol{H}^1(\Gamma)$ and $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and their norms are uniformly bounded by $\|\boldsymbol{g}\|_{1,\Gamma_0}$ and $\|\boldsymbol{f}\|_{0,\widetilde{\Omega}}$. Moreover, from Theorem 2.5 and 2.6 we can conclude that the displacement η is limited, thus the domain Ω can be extended to $\widehat{\Omega}$. The extension of the domain is carried out since the extension of the variables $\boldsymbol{u}$ and p, i.e. $(\widehat{\boldsymbol{u}}, \widehat{p}) \in \boldsymbol{H}^1(\widehat{\Omega}) \times L^2(\widehat{\Omega})$, solve (2.28) over the fixed domain $\widehat{\Omega}$. The restriction of $(\widehat{\boldsymbol{u}}, \widehat{p})$ to Ω gives $(\boldsymbol{u}, p) \in \boldsymbol{H}^{3/2}(\Omega) \times H^{1/2}(\Omega)$.

Schauder's fixed point theorem can be summarized as follows. Let D be a separated topological vector space, $B_R \subset D$ a convex subset, and $T(B_R) \rightarrow B_R$ a continuous function on B_R, equipped with the topology inherited from D. Also let $T(B_R)$ be a compact subset of B_R. Then T has a fixed point, namely, there exists $x \in B_R$ such that $T(x) = x$. Interested reader in Schauder's fixed point theorem can see [82].

Let $(\widehat{\boldsymbol{u}}_1, \widehat{p}_1, \eta_1, \kappa_1) \in \boldsymbol{H}^1(\widehat{\Omega}) \times L_0^2(\widehat{\Omega}) \times H_0^1(I) \times H^1(I)$ be given. Consider now the following mapping

$$T : D = \boldsymbol{H}^1 \times L_0^2 \times H_0^1 \times H^1 \rightarrow A = \boldsymbol{H}^1 \times L_0^2 \times H_0^1 \times H^1,$$

where we rewrite the system of equations (2.28) and (2.24) as

$$
\begin{aligned}
\widehat{\boldsymbol{u}} &= \widehat{\boldsymbol{u}}(\widehat{\boldsymbol{u}}_1, \widehat{p}_1, \eta_1, \kappa_1), & \eta &= \eta(\widehat{\boldsymbol{u}}_1, \widehat{p}_1, \eta_1, \kappa_1), \\
\widehat{p} &= \widehat{p}(\widehat{\boldsymbol{u}}_1, \widehat{p}_1, \eta_1, \kappa_1), & \kappa &= \kappa(\widehat{\boldsymbol{u}}_1, \widehat{p}_1, \eta_1, \kappa_1).
\end{aligned}
\tag{2.30}
$$

We endow the product space $H^1 \times L_0^2 \times H_0^1 \times H^1$ with the norm

$$\|(\widehat{\boldsymbol{u}}_1, \widehat{p}_1, \eta_1, \kappa_1)\| = \|\widehat{\boldsymbol{u}}_1\|_1 + \|\widehat{p}_1\| + \|\eta_1\|_1 + \|\kappa_1\|_1.$$

In order to use the fixed point theorem the solutions of each split equation must be uniformly bounded. Given $(\widehat{\boldsymbol{u}}_1, \widehat{p}_1, \eta_1, \kappa_1)$ it is possible to prove by using standard techniques that $\|\widehat{\boldsymbol{u}}\|_{1,\widehat{\Omega}}$ is uniformly bounded on the fixed domain $\widehat{\Omega}$ by the data norms $\|\boldsymbol{g}\|_{1,\Gamma_0}$ and $\|\boldsymbol{f}\|_{0,\widetilde{\Omega}}$ [43]. This is true since $\widehat{\boldsymbol{g}}$ is zero on the moving boundary, its norm does not change for different η_1 and $\widehat{\boldsymbol{f}}$ is zero outside the domain $\widetilde{\Omega}$. Therefore the $\|\widehat{\boldsymbol{u}}\|_1$ and $\|\widehat{p}\|$ are bounded uniformly by the constants C_u and C_p, respectively, for any $(\widehat{\boldsymbol{u}}_1, \widehat{p}_1, \eta_1, \kappa_1) \in H^1 \times L_0^2 \times H_0^1 \times H^1$.

Since $\boldsymbol{\tau}_s \in \boldsymbol{L}^2(\Gamma)$ we can write explicitly

$$\tau_n = -p\boldsymbol{n}_0 + \nu\big(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^T\big) \cdot \boldsymbol{n}_0 \,, \tag{2.31}$$

obtaining

$$\|\tau_n\|_{0,\Gamma} \leq \|p\|_{0,\Gamma} + \|\boldsymbol{g}\|_{1,\Gamma} \,. \tag{2.32}$$

From Theorem 2.5 and 2.6 we can conclude that norms $\|\tau_s\|_{0,\Gamma_1}$, $\|\eta\|_2$, $\|\kappa\|_1$, $\|r\|_2$ are uniformly bounded by the constants K_t, C_η, C_κ and C_r, respectively. Since $H^2 \subset C^{0,1}(I)$ the curve defined by r is Lipschitz-continuous and bounded for any $(\widehat{\boldsymbol{u}}_1, \widehat{p}_1, \eta_1, \kappa_1, r_1) \in D$. Since the second derivative κ of η is in H^1 then η is in $H^3(I)$. The Sobolev embedding theorem implies η in $C^{1,1}$, which is the required regularity for the boundary Γ_s.

Again, using standard techniques [83], it is possible to show that (2.30) is a continuous mapping with respect to its norm, and it is also linear when $(\boldsymbol{u}_1, p_1, \eta_1, \kappa_1)$ is known. Let now R denote the constant $R = C_u + C_p + C_\eta + C_\kappa$ and let B_R be the ball of radius R. Since for all $(\widehat{\boldsymbol{u}}_1, \widehat{p}_1, \eta_1, \kappa_1) \in D$ we have

$$\|(\widehat{\boldsymbol{u}}, \widehat{p}, \eta, \kappa)\| = \|\boldsymbol{u}\|_1 + \|p\| + \|\eta\|_1 + \|\kappa\|_1 \leq C_u + C_p + C_\kappa + C_\eta = R^2$$

therefore

$$T(B_R) \subset B_R \,.$$

The theorem follows from the compactness of B_R, which can be proved again with standard techniques [83, 84]. When we restrict the domain to Ω we have $(\boldsymbol{u}, p) \in \boldsymbol{H}^{3/2}(\Omega) \times H^{1/2}(\Omega)$ by construction. $\qquad\square$

We now introduce briefly some critical problems caused by the use of mixed (Neumann-Dirichlet) boundary conditions. It is well known that, when $(\boldsymbol{u}, p)$ is a solution of the full Dirichlet problem in an arbitrary polyhedron or a solution of the full Neumann problem in an arbitrary Lipschitz graph polyhedron, then $(\boldsymbol{u}, p) \in \boldsymbol{H}^1(\Omega) \times L^2(\Omega)$ [43]. However, in the mixed boundary conditions in an arbitrary polyhedron case (with the Dirichlet and Neumann boundary conditions prescribed arbitrarily on different faces) this is not always true. In fact, we can write only $(\boldsymbol{u}, p) \in \boldsymbol{W}^{2,8/7+\epsilon}(\Omega) \times W^{1,8/7+\epsilon}(\Omega)$ for some positive ϵ depending on the geometry of the domain.

The fact that the regularity is below the required one needs particular care. Following the theory and the examples in [84], we can conclude that for all Neumann and Dirichlet boundary conditions, where there are no boundary condition changing points, the differentiability can be attained. However,

where there are boundary condition changing points, case by case should be evaluated for mixed boundary conditions. We remark that the solution existence depends strongly on the form of the domain and where the change of boundary conditions is forced. Solution existence can be obtained with techniques similar to those proposed for Dirichlet boundary conditions and it is based on the regularity of the Navier-Stokes equations. Basically we need the local regularity $(\boldsymbol{u}, p) \in \boldsymbol{H}^2 \times H^1$ which is valid outside arbitrarily small edge and vertex neighborhoods when the data are sufficiently smooth. By taking into consideration these notations, we can now analyze the optimization problem.

2.2.3 Existence of the optimal solution

Since we are studying a more complex mathematical problem in comparison with the simple one presented in section 2.1, the existence of the optimal solution is not straightforward. Therefore, we introduce now some basic concepts in order to prove the existence of the optimal solution in the studied case. We recall now the definition of the optimization problem in weak form in the stress case. In fact, in the following we do not consider the control of velocity ($\lambda_1 \neq 0$ in (2.22)). However, it is easy to extend all the formulations presented in the following to the optimization problem with controlled velocity.

Problem 2.3. *Find a state-control set* $(\tau_c, \eta) \in L^2(\Gamma_c) \times \boldsymbol{H}^1(\Gamma_c) \times H^2(\Gamma_0)$ *which minimizes the cost functional*

$$\mathcal{J}(\tau_c, \boldsymbol{g}_c, \eta) = \frac{1}{2} \int_{\Omega_d} |\eta - \eta_d|^2 dx + \frac{\lambda}{2} \|\tau_c\|_{L^2(\Gamma_c)}, \tag{2.33}$$

where $\lambda, \lambda_1 \in \mathbb{R}^+$ *are constant, and under the constraints*

$$
\begin{aligned}
\nu a(\boldsymbol{u}, \boldsymbol{\varphi}) &+ c(\boldsymbol{u}; \boldsymbol{u}, \boldsymbol{\varphi}) + b(\boldsymbol{\varphi}, p) + (\tau_c, \varphi_{n_0})_{\Gamma_{nc}} \\
&+ (\tau_n, \varphi_{n_0})_{\Gamma_d \smallsetminus \Gamma_1} + (\tau_n, \varphi_{n_0})_{\Gamma_n \smallsetminus \Gamma_{nc}} \\
&+ (\beta J_1 \eta, \varphi_{\boldsymbol{n}_0})_{\Gamma_1} + (\gamma J_1 \nabla_{t_0} \eta, \nabla_{t_0} \varphi_{n_0})_{\Gamma_1} \\
&+ (\gamma_1 J_1 \nabla_{t_0} \kappa, \nabla_{t_0} \varphi_{n_0})_{\Gamma_1} = (\boldsymbol{f}, \boldsymbol{\varphi})_{\Omega} &&\forall \boldsymbol{\varphi} \in \boldsymbol{H}_{t_0}^1(\Omega),
\end{aligned}
\tag{2.34}
$$

$$b(\boldsymbol{u}, \psi) = 0 \qquad\qquad\qquad\qquad\quad \forall \psi \in L^2(\Omega), \tag{2.35}$$

$$(\boldsymbol{u}, \boldsymbol{s})_{\Gamma} = (\boldsymbol{n}_0 \, g_n, \boldsymbol{s})_{\Gamma} \qquad\qquad\quad \forall \boldsymbol{s} \in \boldsymbol{H}^{-1/2}(\Gamma), \tag{2.36}$$

$$(\kappa, \psi_1) = (\nabla_{t_0} \eta, \nabla_{t_0} \psi_1) = 0 \qquad\quad \forall \psi_1 \in H_0^1(I), \tag{2.37}$$

$$a(r, \phi) = 0 \qquad\qquad \forall \phi \in H_0^1(\Omega), \qquad (2.38)$$

$$(r, \phi_1)_\Gamma = (J_1\eta, \phi_1)_{\Gamma_s} - (\boldsymbol{u} \cdot \boldsymbol{n}, \phi_1)_{\Gamma_s} \qquad \forall \phi_1 \in L^2(\Gamma). \qquad (2.39)$$

In the system (2.34)-(2.39) we can assume that $\tau_n = 0$ on $\Gamma_n \smallsetminus \Gamma_{nc}$. The τ_n is determined by the g_n on the corresponding boundary. We also set $\eta(0) = \eta(L) = 0$, $\kappa(0) = \kappa_0$ and $\kappa(L) = \kappa_L$. It is important to note that $\boldsymbol{g}$ on Γ_1 should match the boundary vanishing velocity.

Since we are considering a case with mixed boundary conditions, we need to respect the regularity of the solution of the problem for the velocity $\boldsymbol{g}$ and therefore we also require that $\boldsymbol{\tau}_s$ belongs to an admissible set. We assume that the projection of $\boldsymbol{u}$ along the tangent is given in all the regions when Neumann boundary conditions are imposed. To control the normal component of $\boldsymbol{\tau_s}$ we require an improved regularity. On Γ with normal $\boldsymbol{n}_0$, we can usually compute the fluid stresses $\boldsymbol{\tau}_s$ as

$$(\boldsymbol{\tau}_s, \boldsymbol{\varphi})_\Gamma = \nu a(\boldsymbol{u}; \boldsymbol{\varphi}) + c(\boldsymbol{u}; \boldsymbol{u}, \boldsymbol{\varphi}) + b(\boldsymbol{\varphi}, p) - (\boldsymbol{f}, \boldsymbol{\varphi})_\Omega \quad \forall \boldsymbol{\varphi} \in \boldsymbol{H}^1(\Omega),$$

and $\tau_n = \boldsymbol{\tau}_s \cdot \boldsymbol{n}_0$ as

$$\begin{aligned}
(\tau_n, \varphi_{\boldsymbol{n}_0})_\Gamma = \nu a(\boldsymbol{u}; \varphi_{\boldsymbol{n}_0}) + c(\boldsymbol{u}; \boldsymbol{u}, \varphi_{\boldsymbol{n}_0}) \\
+ b(\varphi_{\boldsymbol{n}_0}, p) - (\boldsymbol{f}, \varphi_{\boldsymbol{n}_0})_\Omega \qquad \forall \varphi_{\boldsymbol{n}_0} \in H^1(\Omega),
\end{aligned} \qquad (2.40)$$

for all $(\boldsymbol{u}, p) \in \boldsymbol{H}^{m+1}(\Omega) \times H^m(\Omega)$, $m \geq 1$. We recall that $\boldsymbol{\varphi} = \boldsymbol{t}_0\varphi_{\boldsymbol{t}_0} + \boldsymbol{n}_0\varphi_{\boldsymbol{n}_0}$ and $\boldsymbol{\tau}_s = \boldsymbol{t}_0\tau_t + \boldsymbol{n}_0\tau_n$. When $(\boldsymbol{u}, p) \in \boldsymbol{H}^{3/2}(\Omega) \times H^{1/2}(\Omega)$ the normal component of the stress τ_n is in $L^2(\Gamma)$. We require that τ_n is in $L^2(\Gamma)$ and therefore we define the admissible set

$$\begin{aligned}
\mathcal{T}^{ad} = \{\tau_n \in L^2(\Gamma) \,|\, (\boldsymbol{u}, p) \in \boldsymbol{H}^{3/2}(\Omega) \times H^{1/2}(\Omega) \cap L_0^2(\Omega) \\
\text{satisfies } (2.40)\}.
\end{aligned}$$

For the admissible pressure control we force $\tau_c \in \mathcal{T}_c^{ad}$ with

$$\mathcal{T}_c^{ad} = \boldsymbol{L}^2(\Gamma_{nc}) \cap \mathcal{T}^{ad},$$

and the boundary velocity to be $\boldsymbol{g} = 0\boldsymbol{t}_0 + \boldsymbol{n}_0\, g_n$ such that $g_n \in G_{g_m}^{ad}$, where

$$G_{g_m}^{ad} = \{g_n \in H^1(\Gamma) \,|\, \|\boldsymbol{g}\|_{1,\Gamma} \leq g_m\}. \qquad (2.41)$$

In some optimal problems the admissible displacement $\eta \in H_0^1(I)$ is desired to be limited and therefore we require $\eta \in Q_{ad}$ where

$$Q_{\eta_0,\eta_m}^{ad} = \{\eta \in H_0^1(I) \,|\, c_0 \leq \eta \leq d_0\}. \qquad (2.42)$$

Let $\eta_0 = (c_0 + d_0)/2$ and $\eta_m = (d_0 - c_0)/2$ such as we can write (2.42) as $|\eta - \eta_0| \le \eta_m$. We remark also that $\eta \in Q_{ad}$ and $\kappa \in H^1(I)$ imply $\eta \in H^1(\Omega) \cap H^1_0(I)$.

Under all the presented hypotheses, we can now define the admissible set of states and controls

$$
\begin{aligned}
\mathcal{S}_{ad} = \{ &(\boldsymbol{u}, p, g_n, \tau_c, \eta, \kappa, r) \in \boldsymbol{H}^{3/2}(\Omega_\eta) \times H^1(\Omega_\eta) \cap L^2_0(\Omega_\eta) \times G^{ad}_{g_m} \times \mathcal{T}^{ad}_c(\Gamma_{nc}) \\
&\times Q^{ad}_{\eta_0, \eta_m} \times H^1(I) \times H^1_{\Gamma_\eta \setminus \Gamma_s}(\Omega) \text{ such that } \mathcal{J}(\tau_c, \eta) < \infty \text{ and} \\
&(2.34)\text{-}(2.38) \text{ are satisfied} \} .
\end{aligned}
$$

Moreover, we say that $(\tilde{\boldsymbol{u}}, \tilde{p}, \tilde{g}_n, \tilde{\tau}_c, \tilde{\eta}, \tilde{\kappa}, \tilde{r}) \in \mathcal{S}_{ad}$ is an *optimal solution* if there exists $\epsilon > 0$ such that, for a given $\boldsymbol{f} \in \boldsymbol{L}^2(\widetilde{\Omega})$, $g_n \in H^1(\Omega_d)$ and $\tau_n \in L^2(\Gamma_n)$ we have

$$
\begin{aligned}
&\mathcal{J}(\tilde{\tau}_c, \tilde{\eta}) < \mathcal{J}(\tau_c, \eta) \quad \forall\, (\boldsymbol{u}, p, \boldsymbol{g}_c, \tau_c, \eta, \kappa, r) \in \mathcal{S}_{ad} \text{ satisfying} \\
&\|\boldsymbol{u} - \tilde{\boldsymbol{u}}\|_1 + \|p - \tilde{p}\|_0 + \|\eta - \tilde{\eta}\|_3 + \|\kappa - \tilde{\kappa}\|_1 + \|r - \tilde{r}\|_1 < \epsilon .
\end{aligned}
\tag{2.43}
$$

Now we can introduce the fundamental Theorem on the existence of an optimal solution. In particular, we show that the optimal solution of the problem (2.43) exists when the admissible set $\mathcal{S}_{ad}$ is not empty. When only Neumann or only Dirichlet boundary conditions are considered and no boundary condition type changes are present, the global constraint (2.41) simplifies. However, since we are considering mixed boundary conditions, we need conditions in the agreement with the required regularity. Since the optimal solution, which must be obtained through non-linear iterations, is a local minimum it depends on the initial guess. In our specific geometry, the zero solution of Navier-Stokes equation over a rectangular region is in $\mathcal{S}_{ad}$ when no force is considered. In this case, $\mathcal{S}_{ad}$ is not empty and the zero solution can be used as the initial guess.

Theorem 2.8. *Let $\mathcal{S}_{ad}$ be not empty. There exists an optimal solution $(\boldsymbol{u}, p, \tau_c, g_n, \eta, \kappa, r) \in \mathcal{S}_{ad}$.*

Proof. We first sketch the main idea of this proof, which follows standard techniques in literature. Since the functional $\mathcal{J}$ is bounded from below and $\mathcal{S}_{ad}$ is not empty by hypothesis, we can build a sequence $\{(\boldsymbol{u}^{(k)}, p^{(k)}, g_n^{(k)}, \tau_c^{(k)}, \eta^{(k)}, \kappa^{(k)}, r^{(k)})\} \in \mathcal{S}_{ad}$ that tends to the infimum of the set of the functional values when $(\boldsymbol{u}, p, g_n, \tau_c, \eta, \kappa, r) \in \mathcal{S}_{ad}$. The sequence is uniformly bounded with the corresponding norms. We consider first the fluid equations for the

variables $(\boldsymbol{u}^{(k)}, \boldsymbol{g}^{(k)}, p^{(k)}, \tau_c^{(k)})$ and the extension $\widehat{\Omega}$ of the domain Ω that allows us to study the problem (2.28) for $(\widehat{\boldsymbol{u}}^{(k)}, \widehat{p}^{(k)}, \widehat{\boldsymbol{g}}^{(k)})$ and compute $\tau_c^{(k)}$, as in (2.29), with fixed domain theorems. Since the movement of the boundary Γ_1 is defined by the displacement η, which is bounded above and below by c_2 and d_2, we can easily build a fixed domain $\widehat{\Omega}$ such that its boundaries can be defined as $\widehat{\Gamma} \cap \Gamma = \Gamma \smallsetminus \Gamma_1$. The fixed domain is built to contain all the points of Ω for any (η, κ, r).

Now we consider the following procedure to build the extension $(\widehat{\boldsymbol{u}}^{(k)}, \widehat{p}^{(k)}, \widehat{\boldsymbol{f}})$ in $\boldsymbol{H}^1(\widehat{\Omega}) \times L^2(\widehat{\Omega}) \times \boldsymbol{H}^{-1}(\widehat{\Omega})$ from of the solution $(\boldsymbol{u}^{(k)}, p^{(k)}, \boldsymbol{f})$ in $\boldsymbol{H}^{3/2}(\Omega) \times H^{1/2}(\Omega) \times \boldsymbol{L}^2(\Omega)$. We set $\widehat{\boldsymbol{u}}^{(k)} = \boldsymbol{u}^{(k)}$ and $\widehat{p}^{(k)} = p^{(k)}$ over Ω and zero over $\widehat{\Omega} \smallsetminus \Omega$. By using this extension $\widehat{\boldsymbol{f}} = \boldsymbol{f}$. The test functions can be naturally extended so that they are independent of boundary variations. With this setting, the set $(\widehat{\boldsymbol{u}}^{(k)}, \widehat{g}_n^{(k)}, \widehat{p}^{(k)}, \widehat{\tau}_n^{(k)})$ satisfies the Navier-Stokes system with velocity field and test functions that vanish on $\partial\widehat{\Omega} \smallsetminus \Gamma$. Therefore, the system on the extended domain reads

$$
\begin{aligned}
\nu a(\widehat{\boldsymbol{u}}^{(k)}, \widehat{\boldsymbol{\varphi}}) + c(\widehat{\boldsymbol{u}}^{(k)}; \boldsymbol{u}^{(k)}, \widehat{\boldsymbol{\varphi}}) + b(\widehat{\boldsymbol{\varphi}}, \widehat{p}^{(k)}) \\
= (\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{\varphi}})_{\widehat{\Omega}} && \forall \widehat{\boldsymbol{\varphi}} \in \boldsymbol{H}_{\partial\widehat{\Omega}}^1(\widehat{\Omega}), \\
b(\widehat{\boldsymbol{u}}^{(k)}, \widehat{\psi}) = 0 && \forall \psi \in L_0^2(\widehat{\Omega}), \\
(\widehat{\boldsymbol{u}}^{(k)}, \widehat{\boldsymbol{s}})_{\partial\widehat{\Omega}} = (\boldsymbol{n}_0 \widehat{g}_n^{(k)}, \boldsymbol{s})_{\partial\widehat{\Omega}} && \forall \widehat{\boldsymbol{s}} \in \boldsymbol{H}^{-\frac{1}{2}}(\partial\widehat{\Omega}).
\end{aligned}
\tag{2.44}
$$

Given $\boldsymbol{g}^{(k)} \in \boldsymbol{H}^1(\Gamma)$ with $\|\boldsymbol{g}\|_{1,\Gamma} \leq g_m$ we build the zero extension $\widehat{\boldsymbol{g}}^{(k)}$ on $\partial\widehat{\Omega}$. By using standard techniques with test functions on $\boldsymbol{H}_{\partial\widehat{\Omega}}^1(\widehat{\Omega})$ it is possible to show that the norms $\|\widehat{\boldsymbol{u}}^{(k)}\|_{1,\widehat{\Omega}}$ and $\|\widehat{p}^{(k)}\|_{\widehat{\Omega}}$ are uniformly bounded by their data norms $\|\widehat{\boldsymbol{g}}^{(k)}\|_{1,\widehat{\Gamma}} = \|\boldsymbol{g}\|_{1,\Gamma} \leq g_m$ and $\|\boldsymbol{f}\|_{0,\widehat{\Omega}} = \|\boldsymbol{f}\|_{0,\Omega} \leq f_m$. Therefore, boundary stress norm $\|\widehat{\tau}_n^{(k)}\|_0$ and the control norm $\|\tau_c^{(k)}\|_0$, extracted from (2.29), are uniformly bounded by g_m and f_m. The $(\eta^{(k)}, \kappa^{(k)})$ satisfies

$$
\begin{aligned}
(J_1^{-1}\tau_n^{(k)}, \varphi_n) + (\beta\eta^{(k)}, \varphi_n) + (\gamma\nabla\eta^{(k)}, \nabla\varphi_n) \\
+ (\Gamma_s\nabla\kappa^{(k)}, \nabla\varphi_n) && \forall\varphi_n \in H_0^1(I), \\
(\kappa^{(k)}, \psi_1) = (\nabla\eta^{(k)}, \nabla\psi_1) && \forall\psi_1 \in H_0^1(I).
\end{aligned}
\tag{2.45}
$$

It is clear that, by using the energy balance, the norms $\|\eta^{(k)}\|_1$ and $\|\kappa^{(k)}\|_1$ are bounded by the norm $\|\tau_n^{(k)}\|_0$ which is uniformly bounded by a positive real constant. Finally we have also that $\|\eta^{(k)}\|_3^2 \leq \|\eta\|_1^2 + \|\kappa^{(k)}\|_1^2$ are uniformly bounded. The $r^{(k)}$ satisfies

$$
\begin{aligned}
a(r^{(k)}, \phi) = (f_l, \phi) && \forall\phi \in H_0^1(\Omega), \\
(r^{(k)}, d)_{\Gamma_\eta} = (J_1\eta^{(k)}, d)_{\Gamma_s} && \forall d \in L^2(\Gamma_\eta).
\end{aligned}
\tag{2.46}
$$

Again we can extend the domain from Ω_η to $\widehat{\Omega}$ and the fields $r^{(k)} \in H^1_{\Gamma_\eta \setminus \Gamma_s}(\Omega_\eta)$, $f_l \in L^2(\Omega_\eta)$ to $\widehat{l}^{(k)} \in H^1_0(\widehat{\Omega})$, $\widehat{f}_l \in L^2(\Omega_\eta)(\widehat{\Omega})$ to satisfy the extended equation. By using the Calderon theorem and standard arguments we can prove that $\|r^{(k)}\|_{1,\widehat{\Omega}}$ is uniformly bounded by the uniform bound of the norm of η.

Since the functional $\mathcal{J}$ is bounded and $\mathcal{S}_{ad}$ is not empty, it is possible to state that there exists a uniformly bounded minimizing sequence. Such sequence satisfies the problem on the extended domain $\widehat{\Omega}$. We may then extract a convergent sub-sequence, $(\widehat{\boldsymbol{u}}^{k_1}, \widehat{p}^{k_1}, \widehat{g}_n^{k_1}, \tau_c^{k_1}, \eta^{k_1}, \kappa^{k_1}, r^{k_1})$ such that

$$\lim_{k \to \infty} \mathcal{J}(\tau_c^{(k_1)}, \eta^{(k_1)}) = \inf_{(\boldsymbol{u},p,\tau_c,\eta,\kappa,r)\in\mathcal{S}_{ad}} \mathcal{J}(\tau_c, \eta) \,.$$

and

$$
\begin{array}{cccc}
\widehat{\boldsymbol{u}}^{(k_1)} & \to & \widehat{\boldsymbol{u}} & \text{in} & \boldsymbol{H}^1(\widehat{\Omega}) \\
\widehat{p}^{(k_1)} & \to & \widehat{p} & \text{in} & L^2(\widehat{\Omega}) \\
\tau_c^{(k_1)} & \to & \tau_n & \text{in} & H^{-1/2}(\Gamma_{nc}) \\
\eta^{(k_1)} & \to & \eta & \text{in} & Q_{ad} \\
\kappa^{(k_1)} & \to & \kappa & \text{in} & H^1(I) \\
r^{(k_1)} & \to & r & \text{in} & H^1(\widehat{\Omega}) \\
\widehat{\boldsymbol{u}}^{(k_1)} & \to & \widehat{\boldsymbol{u}} & \text{in} & \boldsymbol{L}^2(\widehat{\Omega}) \; (\textit{strongly}) \\
\widehat{\boldsymbol{u}}^{(k_1)}|_{\partial\widehat{\Omega}} & \to & \widehat{\boldsymbol{u}}_{\partial\widehat{\Omega}} & \text{in} & \boldsymbol{L}^2(\partial\widehat{\Omega}) \; (\textit{strongly}) \,,
\end{array}
\tag{2.47}
$$

for some $(\widehat{\boldsymbol{u}}, \widehat{p}, \widehat{g}_n, \tau_c, \eta, \kappa, \widehat{r}) \in \boldsymbol{H}^1(\widehat{\Omega}) \times L^2(\widehat{\Omega}) \times H^{1/2}(\partial\widehat{\Omega}) \times H^{-1/2}(\widehat{\Gamma}_{nc}) \times Q_{ad} \times H^1(I) \times H^1_0(\widehat{\Omega})$. The last convergence result is obtained through the embedding theorem since $\boldsymbol{H}^1(\widehat{\Omega}) \subseteq \boldsymbol{L}^2(\widehat{\Omega})$ and $H^{1/2}(\partial\widehat{\Omega}) \subseteq L^2(\partial\widehat{\Omega})$.

Now, it is easy to show that the solution of all the linear operators involved in the constraints equations converges to the solution of the equation problem. Therefore the nonlinear term $c(\cdot;\cdot,\cdot)$ is the only one that needs greater attention. For such term we can write

$$c(\widehat{\boldsymbol{u}}^{(k_1)}; \widehat{\boldsymbol{u}}^{(k_1)}, \widehat{\boldsymbol{\varphi}}) = \int_{\widehat{\Gamma}} (\widehat{\boldsymbol{u}}^{(k_1)} \cdot \widehat{\boldsymbol{n}}) \widehat{\boldsymbol{u}}^{(k_1)} \cdot \widehat{\boldsymbol{\varphi}} \, dx - \int_{\widehat{\Omega}} (\widehat{\boldsymbol{u}}^{(k_1)} \cdot \nabla) \widehat{\boldsymbol{\varphi}} \cdot \widehat{\boldsymbol{u}}^{(k_1)} \, d\boldsymbol{x} \,,$$

$\forall \widehat{\boldsymbol{\varphi}} \in C^\infty(\overline{\widehat{\Omega}})$. Then, since $\widehat{\boldsymbol{u}}^{(k_1)} \to \widehat{\boldsymbol{u}}$ in $\boldsymbol{L}^2(\widehat{\Omega})$ and $\widehat{g}^{(k_1)}|_\Gamma \to \widehat{g}|$ strongly in $\boldsymbol{L}^2(\widehat{\Gamma})$, we can take the following limit

$$
\begin{aligned}
\lim_{k_1 \to \infty} c(\widehat{\boldsymbol{u}}^{(k-1)}; \widehat{\boldsymbol{u}}^{(k-1)}, \widehat{\boldsymbol{\varphi}}) &= \int_{\widehat{\Gamma}} (\widehat{\boldsymbol{u}} \cdot \widehat{\boldsymbol{n}}) \widehat{\boldsymbol{u}} \cdot \widehat{\boldsymbol{\varphi}} \, dx - \int_{\widehat{\Omega}} (\widehat{\boldsymbol{u}} \cdot \nabla) \widehat{\boldsymbol{\varphi}} \cdot \widehat{\boldsymbol{u}} \, d\boldsymbol{x} \\
&= c(\widehat{\boldsymbol{u}}; \widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) \qquad \forall \widehat{\boldsymbol{\varphi}} \in C^\infty(\overline{\widehat{\Omega}}) \,.
\end{aligned}
$$

Since $C^\infty(\overline{\widehat{\Omega}})$ is dense in $\boldsymbol{H}^1(\widehat{\Omega})$ we have that

$$\lim_{k_1 \to \infty} c(\widehat{\boldsymbol{u}}^{(k_1)}; \widehat{\boldsymbol{u}}^{(k_1)}, \widehat{\boldsymbol{\varphi}}) = c(\widehat{\boldsymbol{u}}; \widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) \quad \forall \widehat{\boldsymbol{\varphi}} \in \boldsymbol{H}^1(\widehat{\Omega}).$$

Finally we observe that Q_{ad} is a convex set and the functional is weakly lower semi-continuous, therefore $(\widehat{\boldsymbol{u}}, \widehat{p}, \widehat{g}_n, \tau_c, \eta, \kappa, r)$ solves the system of equation in $\boldsymbol{H}^1(\widehat{\Omega}) \times L_0^2(\widehat{\Omega}) \times H^{1/2}(\partial\widehat{\Omega}) \times H^{-1/2}(\Gamma) \times Q_{ad} \times H^1(I) \times H^1(\Omega)$. We have also by construction $\|\widehat{\boldsymbol{g}}\|_{1,\partial\widehat{\Omega}} = \|\boldsymbol{g}\|_{1,\Gamma} \le g_m$. Therefore, we can conclude that the solution limit of the sub-sequences solves the problem (2.44)-(2.45) on the extended domain $\widehat{\Omega}$. In order to conclude the proof we restrict the solution to the domain Ω. The restriction $(\boldsymbol{u}, p, g_n, \tau_c, \eta, \kappa, r)$ of the solution $(\widehat{\boldsymbol{u}}, \widehat{p}, \widehat{g}_n, \widehat{\tau}_n, \eta, \kappa, r)$ in $\boldsymbol{H}^1(\widehat{\Omega}) \times L_0^2(\widehat{\Omega}) \times H^{1/2}(\partial\widehat{\Omega}) \times \boldsymbol{L}^2(\partial\widehat{\Omega}) \times Q_{ad} \times H^1(I) \times H^2(I)$ over Ω is by construction into $\boldsymbol{H}^{3/2}(\Omega) \times H^{1/2}(\Omega) \cap L_0^2(\Omega) \times G_{g_m}^{ad} \times \mathcal{T}_c^{ad} \times Q_{ad} \times H^1(I) \times \times H^1_{\Gamma_\eta \setminus \Gamma_s}(\Omega_\eta)$. $\qquad\square$

2.2.4   The first order necessary condition

In this section, we show that the Lagrange multiplier technique is well posed in the considered mathematical case, and can be used to obtain the first-order necessary condition. In particular, we show that the Lagrangian map is strictly differentiable. As mentioned above, the well-posedness of the mathematical problem studied in these sections is not straightforward. Therefore some preliminary concepts are needed to obtain the first order necessary condition. In order to do that, we follow a formulation similar to the shape control problem studied in [73].

We consider two admissible deformed domains characterized by η, $\bar{\eta} \in Q_{ad}$. We also define the field $\boldsymbol{V}$ that transforms $\Gamma(\eta)$ into $\Gamma(\bar{\eta})$. In our specific case, that can be extended to all the rectangular domains (see Figure 2.2), the η variations can be written as $\delta\eta = \bar{\eta} - \eta$ and $\boldsymbol{V} = (\delta\eta, 0)$. We introduce now the variable t in order to parametrize the boundary Γ as $\Gamma_{\eta + t\delta\eta} = \{\boldsymbol{x}_\eta + t\boldsymbol{V}(\boldsymbol{x}_\eta) \,|\, x_\eta \in \Gamma_\eta\}$, $\forall t \in [0, 1]$. We now introduce the concept of Gateaux-differentiability.

Definition 2.2. *Let us consider a generic functional on $\Omega(\eta)$, $\mathcal{K}(\eta) : \mathcal{C}_{ad} \to \mathbb{R}$, where $\mathcal{C}_{ad}$ is the set of $\eta \in \mathcal{R}_{ad}$. The functional $\mathcal{K}(\eta)$ is Gateaux-differentiable at η in the direction $\delta\eta$ if exists $\mathcal{K}' = (D\mathcal{K}/D\eta)\delta\eta \in \mathbb{R}$ such that*

$$\lim_{t \to 0^+} \frac{|\mathcal{K}(\Omega(\eta_t)) - \mathcal{K}(\Omega(\eta)) - t\mathcal{K}'|}{t} = 0. \tag{2.48}$$

This definition can be extended to Banach spaces, therefore we introduce the following definition.

Definition 2.3. *Let Y be a Banach space, then the map $\boldsymbol{u}(\eta) : \mathcal{C}_{ad} \to Y$ is Gateaux-differentiable at η in the direction $\delta\eta$ if exists $\boldsymbol{u}' \in Y(\Omega(\eta))$ such that*

$$\lim_{t \to 0^+} \frac{\|\boldsymbol{u}(\omega(\eta_t)) - \boldsymbol{u}(\Omega(\eta)) - t\boldsymbol{u}'\|_{Y(\Omega(\eta))}}{t} = 0. \tag{2.49}$$

We now introduce the following lemma that allows us to represent the limit as a boundary integral over Γ_η.

Lemma 2.4. *Take $\eta, \bar{\eta} \in \mathcal{C}_{ad}$, $\hat{y}_1 \in W^{1,1}(\mathbb{R}^2)$, $\hat{y}_2 \in W^{2,1}(\mathbb{R}^2)$,*

$$\mathcal{K}_{\Omega(\eta)} = \int_{\Omega(\eta)} \hat{y}_1(\boldsymbol{x})\, d\boldsymbol{x} \qquad and \qquad \mathcal{K}_{\Gamma(\eta)} = \int_{\Gamma(\eta)} \hat{y}_2(x)\, dx.$$

We can define the "shape derivatives" as

$$\frac{D\mathcal{K}_\Omega}{D\eta} \cdot \delta\eta = \int_{\Omega(\eta)} \nabla \cdot (\boldsymbol{V}\hat{y}_1(\boldsymbol{x}))\, d\boldsymbol{x} = \int_{\Gamma_\eta} \hat{y}_1(x)(\boldsymbol{V} \cdot \hat{\boldsymbol{n}})\, dx, \tag{2.50}$$

$$\frac{D\mathcal{K}_\Gamma}{D\eta} \cdot \delta\eta = \int_{\Gamma_\eta} \Big(\frac{\partial \hat{y}_2}{\partial n} + \kappa \hat{y}_2(x)\Big)(\boldsymbol{V} \cdot \hat{\boldsymbol{n}})\, dx, \tag{2.51}$$

where κ is the curvature of Γ_η and $\hat{\boldsymbol{n}}$ the unit vector normal to it.

In the above lemma $\hat{y}_1$ and $\hat{y}_2$ must be defined on $\mathbb{R}^2$ or extended outside $\Omega(\eta)$ and their gradients must be integrable. Moreover, the hypotheses of the presented Lemma provides that $\hat{y}_1 \in W^{1,1}(\mathbb{R}^2)$. To assure such hypotheses, the solution of the Navier-Stokes system shall be regular enough. If such regularity is not verified, the previous Lemma can be extended to a function $\hat{y}_1 \in L^2(\mathbb{R}^2)$. In this case, the functional is weakly differentiable in H^{-2}, and regular solutions of the Navier-Stokes system are not assured.

In order to use the standard strategies for equality constrained minimization problems (see, for example, [85]) we first need to transform the inequality constraint introduced in (2.42) into an equality. We replace $c_0 \leq \eta \leq d_0$ with

$$|\eta - \eta_0|^2 - \eta_m^2 + s_0^2 = 0 \qquad \forall y \in I, \tag{2.52}$$

for some $s_0 \in H^2(I)$, where $\eta_0 = (c_0 + d_0)/2$, and $\eta_m = (d_0 - c_0)/2$. If we consider $(\boldsymbol{u}, p, \eta, q, r)$ as a solution of the optimal control problem, then there exists s_0 that satisfies (2.52). The Lagrange multiplier method proposed in

this work is based on an embedded technique method where the test functions are kept constant all over the extended domain. Therefore we write the Lagrangian on the extended domain $\widehat{\Omega}$ and then choose the solutions that fit in Ω forcing the boundary values as constraints. We remark that the "hat notation" is used for all the functions defined over $\widehat{\Omega}$.

Let $\widehat{\boldsymbol{B}}_{\boldsymbol{1}} = \boldsymbol{H}^{3/2}(\widehat{\Omega}) \times (L_0^2(\widehat{\Omega}) \cap H^{1/2}(\widehat{\Omega})) \times H_0^1(\Gamma_\eta \smallsetminus \Gamma_s) \times L^2(\Gamma_n) \times H^2(I) \times H^1(I) \times H_0^1(\widehat{\Omega}) \times R \times L^2(I)$, $\widehat{\boldsymbol{B}}_{\boldsymbol{2}} = \boldsymbol{H}^{-1}(\widehat{\Omega}) \times L^2(\widehat{\Omega}) \times \boldsymbol{H}^1(\partial\widehat{\Omega}) \times L^2(\widehat{\Gamma}_n) \times L^2(I) \times H^1(I) \times H^{-1}(\widehat{\Omega}) \times R \times L^1(I)$ and $\widehat{\boldsymbol{B}}_{\boldsymbol{3}} = \boldsymbol{H}^{-1}(\widehat{\Omega}) \times L^2(\widehat{\Omega}) \times \boldsymbol{H}^1(\partial\widehat{\Omega}) \times L^2(\widehat{\Gamma}_n) \times L^2(I) \times L^2(\Gamma_s) \times H_0^1(\widehat{\Omega}) \times R \times L^1(I)$. We can now define the nonlinear mapping $M : \widehat{\boldsymbol{B}}_{\boldsymbol{1}} \to \widehat{\boldsymbol{B}}_{\boldsymbol{3}}$ at $\widehat{\boldsymbol{z}}_0 = (\widehat{\boldsymbol{u}}, \widehat{p}, g_n, \tau_n, \eta, \kappa, \widehat{r}, s_2, s_0)$ by $M(\widehat{\boldsymbol{z}}_0) = \widehat{\boldsymbol{b}}$ with $\widehat{\boldsymbol{b}} = (\widehat{\boldsymbol{l}}_1, \widehat{l}_2, \widehat{\boldsymbol{l}}_3, l_4, l_5, \boldsymbol{l_6}, \widehat{\boldsymbol{l}}_7, l_8, l_9)$ if and only if

$$
\begin{aligned}
&\nu a(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + c(\widehat{\boldsymbol{u}}; \widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + b(\widehat{\boldsymbol{\varphi}}, \widehat{p}) - (\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{\varphi}}) \\
&\quad + (\beta J_1 \eta, \widehat{\varphi}_{n_0})_{\Gamma_1} + (\gamma J_1 \nabla_{t_0} \eta, \nabla_{t_0} + \widehat{\varphi}_{n_0})_{\Gamma_1} \\
&\quad + (\gamma_1 J_1 \nabla_{t_0} \kappa, \nabla_{t_0} \widehat{\varphi}_{n_0})_{\Gamma_1} = (\widehat{\boldsymbol{l}}_1, \widehat{\boldsymbol{\varphi}}) && \forall \widehat{\boldsymbol{\varphi}} \in \boldsymbol{H}_0^1(\widehat{\Omega}) \,, \\
&b(\widehat{\boldsymbol{u}}, \widehat{\psi}) = (\widehat{l}_2, \widehat{\psi}) && \forall \widehat{\psi} \in L^2(\widehat{\Omega}) \,, \\
&(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{s}})_{\partial\widehat{\Omega}} - (\boldsymbol{n}_0 g_n, \widehat{\boldsymbol{s}})_{\Gamma \smallsetminus \Gamma_1} = (\widehat{\boldsymbol{l}}_3, \widehat{\boldsymbol{s}})_{\partial\widehat{\Omega}} && \forall \widehat{\boldsymbol{s}} \in \boldsymbol{H}^{-\frac{1}{2}}(\partial\widehat{\Omega}) \,, \\
&(\boldsymbol{n}_0 \tau_n, \widehat{\boldsymbol{\varphi}})_{\Gamma_n} + \nu a(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + c(\widehat{\boldsymbol{u}}; \widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + b(\widehat{\boldsymbol{\varphi}}, \widehat{p}) \\
&\quad + (\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{\varphi}}) = (\boldsymbol{n}_0 l_4, \widehat{\boldsymbol{\varphi}})_{\Gamma_n} && \forall \widehat{\boldsymbol{\varphi}} \in \boldsymbol{H}_{\boldsymbol{t}_0}^1(\widehat{\Omega}) \,, \\
&(\kappa, \psi_1) + (\nabla_{t_0} \eta, \nabla_{t_0} \psi_1) = (l_5, \psi_1) && \forall \psi_1 \in H_0^1(I) \,, \\
&(\widehat{\boldsymbol{u}}, \boldsymbol{\phi}_1)_{\Gamma_s} - (\boldsymbol{n}_0 J_1 (\widehat{l} - \eta), \boldsymbol{\phi}_1)_{\Gamma_s} = (\boldsymbol{l_6}, \boldsymbol{\phi}_1)_{\Gamma_s} && \forall \boldsymbol{s}_1 \in \boldsymbol{L}^2(\Gamma_s) \,, \\
&a(\widehat{l}, \phi) = (\widehat{l}_7, \phi) && \forall \phi \in H_0^1(\widehat{\Omega}) \\
&\|g_n\|_{1, \partial\widehat{\Omega} \cap \Gamma_\eta}^2 - g_m^2 + \widehat{s}_2^2 = l_8 \,, \\
&|\eta - \eta_0|^2 - \eta_m^2 + s_0^2 = l_9 && \forall t \in I \,,
\end{aligned}
\tag{2.53}
$$

We recall that τ_c is obtained when τ_n is restricted to Γ_c. From the definition of $(\widehat{\boldsymbol{l}}_1, \widehat{l}_2, \widehat{\boldsymbol{l}}_3, l_4, l_5, \boldsymbol{l_6}, \widehat{\boldsymbol{l}}_7, l_8, l_9)$, we can state that the set of constraint equations in the optimal shape control problem can be expressed as $M(\widehat{\boldsymbol{z}}_0) = (\boldsymbol{0}, 0, \boldsymbol{0}, 0, 0, \boldsymbol{0}, 0, 0, 0)$, where $\widehat{\boldsymbol{z}}_0 = (\widehat{\boldsymbol{u}}, \widehat{p}, \boldsymbol{g}_n, \tau_n, \eta, \kappa, \widehat{r}, s_0, s_2)$ is the optimal solution. Note that the boundary velocity over the fixed boundary is set in the third equation while the boundary velocity over the solid is included in the sixth equation. In particular, when $\boldsymbol{l}_6 = 0$ the tangential component $\boldsymbol{u} \cdot \boldsymbol{t}_0 = g_t$ is set to zero and the normal one is zero only when $r = r_0 + \eta$ and $\boldsymbol{u} \cdot \boldsymbol{n}_0 = g_n = 0$. It is very useful to define a fictitious velocity to enable an iterative algorithm to force the boundary on the correct displacement imposed by the solid model: when $\widehat{r} = \eta$ the velocity $\boldsymbol{u} \cdot \boldsymbol{n}_0$ must vanish. The

fourth equation forces Neumann boundary conditions and control in the set $\mathcal{T}^{ad}$.

Given $\widehat{\boldsymbol{z}}_1 = (\widehat{\boldsymbol{u}}_1, \widehat{p}_1, g_{1n}\tau_{1n}, \eta_1, \kappa_1, \widehat{r}_1, s_{21}, s_{01}) \in \mathcal{S}_{ad}$ we can now define the nonlinear mapping $Q : \widehat{\boldsymbol{B}_1} \to \mathbb{R} \times \widehat{\boldsymbol{B}_3}$. For $a \in \mathbb{R}$ we set $Q(\widehat{\boldsymbol{z}}_0) = (a, \widehat{\boldsymbol{b}})$ if and only if

$$
Q = \begin{pmatrix} \mathcal{J}(\tau_n, \eta) - \mathcal{J}(\tau_{n1}, \eta_1) \\ M(\widehat{\boldsymbol{z}}_0) \end{pmatrix} = \begin{pmatrix} a \\ (\widehat{\boldsymbol{l}}_1, \widehat{l}_2, \widehat{\boldsymbol{l}}_3, l_4, l_5, \boldsymbol{l_6}, l_7, l_8) \end{pmatrix}. \quad (2.54)
$$

Mappings differentiability

In this paragraph, we rely on the definition of strict differentiability in order to show that the above mappings M and Q are strictly differentiable. We recall now the definition of strict differentiability (see [85]).

Definition 2.4. *Let X and Y be Banach spaces. The mapping $\phi : X \to Y$ is strictly differentiable at $x \in X$ if there exist a bounded, linear mapping $D : X \to Y$ so that $\forall \varepsilon > 0$, $\exists \delta > 0$ ensuring that for $x_1, x_2 \in X$, $\|x - x_1\|_X < \delta$ and $\|x - x_2\|_X < \delta$, then*

$$
\|\phi(x_1) - \phi(x_2) - D(x_1 - x_2)\|_Y \leq \varepsilon \|x_1 - x_2\|_X, \quad (2.55)
$$

where the strict derivative is denoted as $D = \phi'(x)$.

In this work we will consider $X = \widehat{\boldsymbol{B}_1}$ and $Y = \widehat{\boldsymbol{B}_2}$. Let $\boldsymbol{z}_0 = (\widehat{\boldsymbol{u}}, \widehat{p}, g_n, \tau_n, \eta, \kappa, r, s_2, s_0)$ be in $\widehat{\boldsymbol{B}_1}$ and consider, as derivative map, the bounded linear operator $M' : \widehat{\boldsymbol{B}_1} \to \widehat{\boldsymbol{B}_2}$, where $M'(\widetilde{\boldsymbol{z}}) \cdot \widetilde{\boldsymbol{z}}_0 = \bar{\boldsymbol{b}}$ for $\widetilde{\boldsymbol{z}}_0 = (\widetilde{\boldsymbol{u}}, \widetilde{p}, \widetilde{g}_n, \widetilde{\tau}_n, \widetilde{\eta}, \widetilde{\eta}, \widetilde{r}, \widetilde{s}_2, \widetilde{s}_0) \in \widehat{\boldsymbol{B}_1}$ and $\bar{\boldsymbol{b}} = (\bar{l}_1, \bar{l}_2, \bar{l}_3, \bar{l}_4, \bar{l}_5, \bar{l}_6, \bar{l}_7, \bar{l}_8) \in \widehat{\boldsymbol{B}_2}$ defined as

$$
\begin{aligned}
\nu a(\widetilde{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) &+ c(\widetilde{\boldsymbol{u}}; \widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + c(\widehat{\boldsymbol{u}}; \widetilde{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + b(\widehat{\boldsymbol{\varphi}}, \widetilde{p}) \\
&+ (\beta J_1 \widetilde{\eta}, \widehat{\varphi}_{n_0})_{\Gamma_1} + (\gamma J_1 \nabla_{t_0} \widetilde{\eta}, \nabla_{t_0} + \widehat{\varphi}_{n_0})_{\Gamma_1} \\
&+ (\gamma_1 J_1 \nabla_{t_0} \widetilde{\kappa}, \nabla_{t_0} \widehat{\varphi}_{n_0})_{\Gamma_1} = (\bar{\boldsymbol{l}}_1, \widehat{\boldsymbol{\varphi}}) && \forall \widehat{\boldsymbol{\varphi}} \in \boldsymbol{H}_0^1(\widehat{\Omega}), \\
b(\widetilde{\boldsymbol{u}}, \widehat{\psi}) &= (\bar{l}_2, \widehat{\psi}) && \forall \widehat{\psi} \in L^2(\widehat{\Omega}), \\
(\widetilde{\boldsymbol{u}}, \widehat{\boldsymbol{s}})_{\partial\widehat{\Omega}} &- (\boldsymbol{n}_0 \widetilde{g}_n, \widehat{\boldsymbol{s}})_{\Gamma \smallsetminus \Gamma_1} = (\bar{\boldsymbol{l}}_3, \widehat{\boldsymbol{s}})_{\partial\widehat{\Omega}} && \forall \widehat{\boldsymbol{s}} \in \boldsymbol{H}^{-1/2}(\partial\widehat{\Omega}), \\
(\boldsymbol{n}_0 \widetilde{\tau}_n, \widehat{\boldsymbol{\varphi}})_{\Gamma_n} &+ \nu a(\widetilde{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + c(\widetilde{\boldsymbol{u}}; \widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) \\
&+ c(\widehat{\boldsymbol{u}}; \widetilde{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + b(\widehat{\boldsymbol{\varphi}}, \widetilde{p}) = (\boldsymbol{n}_0 \widetilde{l}_4, \widehat{\boldsymbol{\varphi}})_{\Gamma_n} && \forall \widehat{\boldsymbol{\varphi}} \in \boldsymbol{H}_{t_0}^1(\widehat{\Omega}), \quad (2.56) \\
(\widetilde{\kappa}, \psi_1) &- (\nabla_{t_0} \widetilde{\eta}, \nabla_{t_0} \psi_1) = (\widetilde{l}_5, \psi_1) && \forall \psi_1 \in H_0^1(I),
\end{aligned}
$$

$$(\widetilde{\boldsymbol{u}}, \boldsymbol{s}_1) + \int_{\Gamma_s} \left(\boldsymbol{V}(\widetilde{\eta}) \cdot \hat{\boldsymbol{n}}\right)\left(\chi + \frac{\partial}{\partial \hat{\boldsymbol{n}}}\right)\widehat{\boldsymbol{u}} \cdot \boldsymbol{s}_1 \, d\boldsymbol{x}$$

$$- (\boldsymbol{n}_0 J_1(\widetilde{r} - \widetilde{\eta}), \boldsymbol{s}_1) = (\bar{\boldsymbol{l}}_6, \boldsymbol{s}_1) \qquad\qquad \forall \boldsymbol{s}_1 \in \boldsymbol{L}^2(\Gamma_s),$$

$$a(\widetilde{r}, \widehat{\phi}) = (\bar{l}_7, \widehat{\phi}) \qquad\qquad \forall \widehat{\phi} \in H_0^1(\widehat{\Omega}),$$

$$\int_{\partial\widehat{\Omega}\cap\Gamma_\eta} 2(g_n' \widetilde{g}_n' + g_n \widetilde{g}_n) \, ds + 2s_2 \widetilde{s}_2 = \bar{l}_8,$$

$$2(\eta - \eta_0)\widetilde{\eta} + 2s_0 \widetilde{s}_0 = \bar{l}_9,$$

with $\boldsymbol{V}(\widetilde{\eta}) = (\widetilde{\eta}, 0)$. The normal vector to Γ_1 is denoted by $\hat{\boldsymbol{n}}$ and the boundary curvature is χ. By $\widetilde{\boldsymbol{g}}_n'$ we denote the derivative of $\widetilde{\boldsymbol{g}}_n$.

We introduce now a Lemma on the differentiability of the mappings M and Q. We define the differentiability operators M' and Q', respectively.

Lemma 2.5. *Let the nonlinear mapping $M : \widehat{\boldsymbol{B}}_1 \to \widehat{\boldsymbol{B}}_3$ be defined by (2.53) and $Q : \widehat{\boldsymbol{B}}_1 \to \mathbb{R} \times \widehat{\boldsymbol{B}}_3$ by (2.54) defined at a point $\widehat{\boldsymbol{z}}_0 = (\widehat{\boldsymbol{u}}, \widehat{p}, g_n, \tau_n, \eta, \kappa, \widehat{r}, s_2, s_0) \in \widehat{\boldsymbol{B}}_1$. We consider also the bounded linear operator $M' : \widehat{\boldsymbol{B}}_1 \to \widehat{\boldsymbol{B}}_2$, where $M'(\widehat{\boldsymbol{z}}_0) \cdot \widetilde{\boldsymbol{z}}_0 = \bar{\boldsymbol{b}}$ for $\widetilde{\boldsymbol{z}}_0 = (\widetilde{\boldsymbol{u}}, \widetilde{p}, \widetilde{g}_n, \widetilde{\tau}_n, \widetilde{\eta}, \widetilde{\eta}, \widetilde{r}, \widetilde{s}_2, \widetilde{s}_0) \in \widehat{\boldsymbol{B}}_1$ and $\bar{\boldsymbol{b}} = (\bar{l}_1, \bar{l}_2, \bar{l}_3, \bar{l}_4, \bar{l}_5, \bar{l}_6, \bar{l}_7, \bar{l}_8, \bar{l}_9) \in \widehat{\boldsymbol{B}}_2$. Then, the mappings M and Q are strictly differentiable at the point $\widehat{\boldsymbol{z}}_0 \in \widehat{\boldsymbol{B}}_1$ and its strict derivative is given by M' in (2.56).*

Consider the nonlinear operator $Q'(\widehat{\boldsymbol{z}}_0) : \widehat{\boldsymbol{B}}_1 \to \mathbb{R} \times \widehat{\boldsymbol{B}}_2$, where $Q'(\widehat{\boldsymbol{z}}_0) \cdot \widetilde{\boldsymbol{z}}_0 = (\bar{a}, \bar{\boldsymbol{b}})$ for $\bar{a} \in R$. If we set

$$\mathcal{J}'(\tau_c, \eta) \cdot (\widetilde{\boldsymbol{z}}_0) = \int_{\Omega_d} \widetilde{\eta}(\eta - \eta_d) dx + \lambda \int_{\Gamma_{nc}} \widetilde{\tau}_c \, \tau_c dx, \qquad (2.57)$$

then the strict derivative of Q at a point $\widehat{\boldsymbol{z}}_0$ is given by Q' if and only if

$$\begin{pmatrix} \mathcal{J}'(\tau_c, \eta) \cdot \widetilde{\boldsymbol{z}}_0 \\ M'(\widehat{\boldsymbol{z}}_0) \cdot \widetilde{\boldsymbol{z}}_0 \end{pmatrix} = \begin{pmatrix} \widetilde{a} \\ \bar{\boldsymbol{b}} \end{pmatrix}. \qquad (2.58)$$

Proof. Since the operators $a(\cdot, \cdot)$, $b(\cdot, \cdot)$ and $c(\cdot; \cdot, \cdot)$ are continuous we have that $M'(\boldsymbol{z}_0)$ is bounded. The linearity of such operator is straightforward to demonstrate. For the same reasons we can assume that $Q'(\boldsymbol{z}_0)$ is bounded and linear. Therefore for all $\boldsymbol{z}_0 \in (\widehat{\boldsymbol{u}}, \widehat{p}, g_n, \tau_c, \eta, q, \widehat{r}, s_0, s_2) \in \widehat{\boldsymbol{B}}_1$ and for all $\widehat{\boldsymbol{z}}_a = (\widehat{\boldsymbol{\omega}}_a, \widehat{p}_a, \boldsymbol{\zeta}, \boldsymbol{\theta}, \mu_2, \boldsymbol{g}_a, l_a, s_{2a}, s_{0a}) \in \widehat{\boldsymbol{B}}_2^*$, we have

$$\left\langle \boldsymbol{z}_a, \left(M(\boldsymbol{z}_1) - M(\boldsymbol{z}_2) - M'(\boldsymbol{z}_0) \cdot (\boldsymbol{z}_1 - \boldsymbol{z}_2)\right)\right\rangle =$$

$$= c(\widehat{\boldsymbol{u}}_1; \widehat{\boldsymbol{u}}_1, \widehat{\boldsymbol{\omega}}) - c(\widehat{\boldsymbol{u}}_2; \widehat{\boldsymbol{u}}_2, \widehat{\boldsymbol{\omega}}) - \left(c(\widehat{\boldsymbol{u}}_1 - \widehat{\boldsymbol{u}}_2; \widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\omega}}) + c(\widehat{\boldsymbol{u}}; \widehat{\boldsymbol{u}}_1 - \widehat{\boldsymbol{u}}_2, \widehat{\boldsymbol{\omega}})\right)$$

$$+ c(\widehat{\boldsymbol{u}}_1; \widehat{\boldsymbol{u}}_1, \widehat{\boldsymbol{\theta}}) - c(\widehat{\boldsymbol{u}}_2; \widehat{\boldsymbol{u}}_2, \widehat{\boldsymbol{\theta}}) - \big(c(\widehat{\boldsymbol{u}}_1 - \widehat{\boldsymbol{u}}_2; \widehat{\boldsymbol{u}}, \widehat{\boldsymbol{\theta}}) + c(\widehat{\boldsymbol{u}}; \widehat{\boldsymbol{u}}_1 - \widehat{\boldsymbol{u}}_2, \widehat{\boldsymbol{\theta}})\big)$$

$$+ \int_{\Gamma_{11}} \boldsymbol{\zeta} \cdot \boldsymbol{u}_1 dx - \int_{\Gamma_{12}} \boldsymbol{\zeta} \cdot \boldsymbol{u}_2 dx - \int_{\Gamma_1} \boldsymbol{\zeta} \cdot (\boldsymbol{u}_1 - \boldsymbol{u}_2) dx$$

$$- \int_{\Gamma_1} \left(\chi + \frac{\partial}{\partial \widehat{\boldsymbol{n}}}\right) \boldsymbol{\zeta} \cdot \boldsymbol{u} \big(\boldsymbol{V}(\eta_1 - \eta_2) \cdot \widehat{\boldsymbol{n}}\big) \, dx + \tau_2\big(s_{21}^2 - s_{22}^2 - 2s_2(s_{21} - s_{22})\big)$$

$$+ \tau_2 \int_{\partial \widehat{\Omega} \cap \Gamma} (g_{1n}^{'2} + g_{1n}^2) - (g_{2n}^{'2} + g_{2n}^2) - 2g_n'(g_{1n}' - g_{2n}') - 2g_n(g_{1n} - g_{2n}) \, ds$$

$$+ \int_I \tau_0(\eta_1 - \eta_0)^2 + -\tau_0(\eta_2 - \eta_0)^2 - 2\tau_0(\eta - \eta_0)(\eta_1 - \eta_2)$$

$$+ \int_I \tau_0 s_{01}^2 - \tau_0 s_{02}^2 - 2\tau_0 s_2(s_{01} - s_{02}) \, .$$

We can cancel out all the linear terms, e.g. $a(\widehat{\boldsymbol{u}}_1, \widehat{\boldsymbol{\omega}}) - a(\widehat{\boldsymbol{u}}_2, \widehat{\boldsymbol{\omega}}) - a(\widehat{\boldsymbol{u}}_1 - \widehat{\boldsymbol{u}}_2, \widehat{\boldsymbol{\omega}}) = 0$. The terms with J_1 can be computed over the reference interval I and therefore do not have shape derivative.

From the Lemma 2.4 we have that $\forall \varepsilon_1 > 0$, $\exists \delta_1 > 0$ such that

$$\left| \int_{\Gamma_{1i}} \boldsymbol{\zeta} \cdot \boldsymbol{u}_i dx - \int_{\Gamma_1} \boldsymbol{\zeta} \cdot \boldsymbol{u}_i dx - t_1 \int_{\Gamma_1} \left(\chi + \frac{\partial}{\partial \widehat{\boldsymbol{n}}}\right) \boldsymbol{\zeta} \cdot \boldsymbol{u}_i \big(\boldsymbol{V}(\eta_i - \eta) \cdot \widehat{\boldsymbol{n}}\big) \, dx \right|$$
$$< \frac{\varepsilon_1}{4} \|\eta_1 - \eta_2\|_{H^2(I)} \, ,$$
$$(2.59)$$

for $i = 1, 2$ and $t_1 \leq \delta_1$. By using (2.59) we have

$$\left| \left\langle \boldsymbol{z}_a, \Big(M(\boldsymbol{z}_1) - M(\boldsymbol{z}_2) - M'(\boldsymbol{z}_0) \cdot (\boldsymbol{z}_1 - \boldsymbol{z}_2)\Big) \right\rangle \right| \leq \frac{\varepsilon}{2} \|\eta_1 - \eta_2\|_{H^2(I)} +$$

$$+ |c(\boldsymbol{u}_1 - \boldsymbol{u}_2; \boldsymbol{u} - \boldsymbol{u}_1, \widehat{\boldsymbol{\omega}}) + c(\boldsymbol{u} - \boldsymbol{u}_2; \boldsymbol{u}_1 - \boldsymbol{u}_2, \widehat{\boldsymbol{\omega}})|$$

$$+ |c(\boldsymbol{u}_1 - \boldsymbol{u}_2; \boldsymbol{u} - \boldsymbol{u}_1, \widehat{\boldsymbol{\theta}}) + c(\boldsymbol{u} - \boldsymbol{u}_2; \boldsymbol{u}_1 - \boldsymbol{u}_2, \widehat{\boldsymbol{\theta}})|$$

$$+ \tau_2 \int_I \big((g_{n1}' - g_{n2}')(g_{n1}' - g_n' + g_{n2}' - g_n')$$

$$+ (g_{n1} - g_{n2})(g_{n1} - g_n + g_{n2} - g_n)\big) \, dx$$

$$+ \tau_2 \int_I (s_{21} - s_{22})(s_{21} - s_2 + s_{22} - s_2) dx$$

$$+ \tau_0 \int_I (s_{01} - s_{02})(s_{01} - s_0 + s_{02} - s_0) \, dx$$

$$+ \tau_0 \int_I \big((\eta_1 - \eta_2)(\eta_1 - \eta + \eta_2 - \eta)$$

$$+ (s_{01} - s_{02})(s_{01} - s_0 + s_{02} - s_0)\big) \, dx \, ,$$
$$(2.60)$$

By using the Sobolev embedding theorem, the trace theorem, and the continuity of the operator $c(\cdot; \cdot, \cdot)$ it is possible to bound the norm of the above

scalar product. Therefore, we can state that $\exists\, C_1, C_2, C_3 \in \mathbb{R}^+$ such that

$$
\begin{aligned}
&\left\| M(\boldsymbol{z}_1) - M(\boldsymbol{z}_2) - M'(\boldsymbol{z}_0) \cdot (\boldsymbol{z}_1 - \boldsymbol{z}_2) \right\|_{\widehat{\boldsymbol{B}_2}} \\
&\leq \frac{\varepsilon}{2} \|\eta_1 - \eta_2\|_{H^2(I)} + C_1 \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_1 (\|\boldsymbol{u} - \boldsymbol{u}_1\|_1 + \|\boldsymbol{u} - \boldsymbol{u}_2\|_1) \\
&\quad + C_2 \big(\|g'_{n1} - g'_{n2}\| \, (\|g'_{n1} - g'_n\| + \|g'_{n2} - g'_n\|) \\
&\qquad + \|g_{n1} - g_{n2}\| (\|g_{n1} - g_n\| + \|g_{n2} - g_n\|) \\
&\qquad + \|s_{21} - s_{22}\| (\|s_2 - s_{21}\| + \|s_2 - s_{22}\|) \big) \\
&\quad + C_3 \big(\|\eta_1 - \eta_2\|_2 (\|\eta_1 - \eta\|_2 + \|\eta_1 - \eta_2\|_2) \\
&\qquad + \|s_{01} - s_{02}\|_2 (\|s_0 - s_{01}\|_2 + \|s_0 - s_{02}\|_2) \big) .
\end{aligned}
$$

For the norm properties there exists a positive constant C_4 such that

$$
\begin{aligned}
&\left\| M(\boldsymbol{z}_1) - M(\boldsymbol{z}_2) - M'(\boldsymbol{z}_0) \cdot (\boldsymbol{z}_1 - \boldsymbol{z}_2) \right\|_{\widehat{\boldsymbol{B}_2}} \\
&\leq \frac{\varepsilon}{2} \|\eta_1 - \eta_2\|_{H^2(I)} + C_4 \left\| \boldsymbol{z}_1 - \boldsymbol{z}_2 \right\|_{\widehat{\boldsymbol{B}_1}} \\
&\quad \left(\left\| (\widehat{\boldsymbol{u}} - \widehat{\boldsymbol{u}}_1, \widehat{p} - \widehat{p}_1, p_c - p_{c1}, \eta - \eta_1, r - r_1, s_0 - s_{01}) \right\|_{\widehat{\boldsymbol{B}_1}} \right. \\
&\qquad \left. + \left\| (\widehat{\boldsymbol{u}} - \widehat{\boldsymbol{u}}_2, \widehat{p} - \widehat{p}_2, p_c - p_{c2}, \eta - \eta_2, r - r_2, s_0 - s_{02}) \right\|_{\widehat{\boldsymbol{B}_1}} \right).
\end{aligned}
\tag{2.61}
$$

Now $\forall \varepsilon > 0$, by choosing $\delta = \dfrac{\varepsilon}{4 C_4}$ we have that, whenever $\left\| \boldsymbol{z} - \boldsymbol{z}_1 \right\|_{\widehat{\boldsymbol{B}_1}} < \delta$ and $\left\| \boldsymbol{z} - \boldsymbol{z}_2 \right\|_{\widehat{\boldsymbol{B}_1}} < \delta$ we obtain

$$
\begin{aligned}
&\left\| M(\boldsymbol{z}_1) - M(\boldsymbol{z}_2) - M'(\boldsymbol{z}_0) \cdot (\boldsymbol{z}_1 - \boldsymbol{z}_2) \right\|_{\widehat{\boldsymbol{B}_2}} \\
&\leq \varepsilon \left\| (\widehat{\boldsymbol{u}}_1 - \widehat{\boldsymbol{u}}_2, \widehat{p}_1 - \widehat{p}_2, p_{c1} - p_{c2}, \eta_1 - \eta_2, q_1 - q_2, \right. \\
&\qquad\qquad \left. r_1 - r_2, s_{01} - s_{02}) \right\|_{\widehat{\boldsymbol{B}_1}}.
\end{aligned}
\tag{2.62}
$$

Therefore, the mapping M is strictly differentiable on $\widehat{\boldsymbol{B}_1}$ and its strict derivative is M'. By using (2.62) and Lemma 2.4, it is easy to demonstrate that also Q is strictly differentiable on $\widehat{\boldsymbol{B}_1}$ and its strict derivative is Q'. $\qquad\square$

As anticipated, we now restrict all the variables to the domain Ω. In particular, the solution $(\boldsymbol{u}, p)$ can be seen as the restriction to the domain Ω of the solution $(\widehat{\boldsymbol{u}}, \widehat{p})$ of the Navier-Stokes system over $\widehat{\Omega}$. Using the same approach, we call $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ the restrictions of $\widehat{\boldsymbol{B}_1}$ and $\widehat{\boldsymbol{B}_2}$ from $\widehat{\Omega}$ to Ω.

With this notation, we can restrict the mapping M' from (2.56) to

$$
\begin{aligned}
&\nu a(\widetilde{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + c(\widetilde{\boldsymbol{u}}; \boldsymbol{u}, \widehat{\boldsymbol{\varphi}}) + c(\boldsymbol{u}; \widetilde{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + b(\widehat{\boldsymbol{\varphi}}, \widetilde{p}) \\
&\qquad + (\beta J_1 \widetilde{\eta}, \widehat{\varphi}_{n_0})_{\Gamma_1} + (\gamma J_1 \nabla_{t_0}\widetilde{\eta}, \nabla_{t_0} + \widehat{\varphi}_{n_0})_{\Gamma_1} \\
&\qquad + (\gamma_1 J_1 \nabla_{t_0}\widetilde{\kappa}, \nabla_{t_0}\widehat{\varphi}_{n_0})_{\Gamma_1} = (\bar{\boldsymbol{l}}_1, \widehat{\boldsymbol{\varphi}}) && \forall \widehat{\boldsymbol{\varphi}} \in \boldsymbol{H}^1_{\Gamma^*}(\Omega)\,, \\
&b(\widetilde{\boldsymbol{u}}, \widehat{\psi}) = (\bar{l}_2, \widehat{\psi}) && \forall \widehat{\psi} \in L^2(\Omega)\,, \\
&(\widetilde{\boldsymbol{u}}, \widehat{\boldsymbol{s}})_{\Gamma^*} - (\boldsymbol{n}_0 \widetilde{g}_n, \widehat{\boldsymbol{s}})_{\Gamma^*} = (\bar{\boldsymbol{l}}_3, \widehat{\boldsymbol{s}})_{\Gamma^*} && \forall \widehat{\boldsymbol{s}} \in \boldsymbol{H}^{-\frac{1}{2}}(\Gamma^*)\,, \\
&(\boldsymbol{n}_0 \widetilde{\tau}_n, \widehat{\boldsymbol{\varphi}})_{\Gamma_n} + \nu a(\widetilde{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + c(\widetilde{\boldsymbol{u}}; \boldsymbol{u}, \widehat{\boldsymbol{\varphi}}) + c(\boldsymbol{u}; \widetilde{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) \\
&\qquad + b(\widehat{\boldsymbol{\varphi}}, \widetilde{p}) = (\boldsymbol{n}_0 \bar{l}_4, \widehat{\boldsymbol{\varphi}})_{\Gamma_n} && \forall \widehat{\boldsymbol{\varphi}} \in \boldsymbol{H}^1_{t_0}(\Omega)\,, \\
&(\widetilde{\kappa}, \psi_1) - (\nabla_{t_0}\widetilde{\eta}, \nabla_{t_0}\psi_1) = (\bar{l}_5, \psi_1) && \forall \psi_1 \in H^1_0(I)\,, \\
&(\widetilde{\boldsymbol{u}}, \boldsymbol{s}_1)_{\Gamma_s} + \int_{\Gamma_s} \big(\boldsymbol{V}(\widetilde{\eta}) \cdot \widehat{\boldsymbol{n}}\big)\Big(\chi + \frac{\partial}{\partial \boldsymbol{n}}\Big)\widehat{\boldsymbol{u}} \cdot \boldsymbol{s}_1 \, d\boldsymbol{x} \\
&\qquad + (\boldsymbol{n}_0 J_1(\widetilde{r} - \widetilde{\eta}), \boldsymbol{s}_1)_{\Gamma_s} = (\bar{\boldsymbol{l}}_6, \boldsymbol{s}_1)_{\Gamma_s} && \forall \boldsymbol{s}_1 \in \boldsymbol{L}^2(\Gamma_s)\,, \\
&a(l, \phi) = (\bar{l}_7, \phi) && \forall \phi \in H^1_0(\Omega) \\
&\int_{\Gamma_\eta} 2(g'_n \widetilde{g}'_n + g_n \widetilde{g}_n)\, ds + 2s_2 \widetilde{s}_2 \ = \ \bar{l}_8\,, \\
&2(\eta - \eta_0)\widetilde{\eta} + 2s_0 \widetilde{s}_0 \ = \ \bar{l}_9\,,
\end{aligned}
\tag{2.63}
$$

where $\Gamma^* = \Gamma \smallsetminus \Gamma_1$. It is worth noticing that the map $M'(\boldsymbol{u}, p, g_n, \tau_n, \eta, \kappa, r, s_0, s_2)$ assumes the same values of the restriction of $M'(\widehat{\boldsymbol{u}}, \widehat{p}, g_n, \tau_n, \eta, \kappa, r, s_0, s_2)$ to Ω since no further information is necessary to identify the function over the restricted domain. For this reason the maps (2.56) and (2.63) are equivalent on Ω. The same statements holds for $Q'(\boldsymbol{u}, p, g_n, \tau_n, \eta, \kappa, r, s_0, s_2)$.

Now we introduce some additional properties of the derivatives of the mappings M and Q. To do so, we formulate the following Lemma.

Lemma 2.6. *Let $\boldsymbol{z}_0 = (\boldsymbol{u}, p, g_n, \tau_n, \eta, \kappa, r, s_0, s_2) \in \boldsymbol{B}_1$ be a solution of the optimal control problem. We state that*

1. *the operator $M'(\boldsymbol{z}_0)$ has closed range in $\boldsymbol{B}_2$,*

2. *the operator $Q'(\boldsymbol{z}_0)$ has closed range in $\mathbb{R} \times \boldsymbol{B}_2$,*

3. *the operator $Q'(\boldsymbol{z}_0)$ is not onto in $\mathbb{R} \times \boldsymbol{B}_2$.*

Proof. (1.) The first statement can be proved by showing that the range of all the equations of (2.63) is closed. In particular, we consider again the split system. Now, we have to show that the range of the fluid equations is closed,

therefore we focus on the fluid problem (2.63) by changing the test functions from $\widehat{\boldsymbol{\varphi}} \in \boldsymbol{H}^1_{\Gamma \smallsetminus \Gamma_1}(\Omega)$ to $\widehat{\boldsymbol{\varphi}} \in \boldsymbol{H}^1_0(\Omega)$. Given $\boldsymbol{z}_0 = (\boldsymbol{u}, p, g_n, \tau_n, \eta, \kappa, r, s_0, s_2) \in \boldsymbol{B_1}$ and $(\bar{\boldsymbol{l}}_1, \bar{\boldsymbol{l}}_2, \bar{\boldsymbol{l}}_3, \bar{\boldsymbol{l}}_6, \bar{\boldsymbol{l}}_8) \in \boldsymbol{H}^1_0(\Omega_\eta) \times L^2(\Omega_\eta) \times \boldsymbol{H}^1_0(\Gamma \smallsetminus \Gamma_s) \times \boldsymbol{L}^2(\Gamma_s) \times R$ we have

$$
\begin{aligned}
\nu a(\widetilde{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) &+ c(\widetilde{\boldsymbol{u}}; \boldsymbol{u}, \widehat{\boldsymbol{\varphi}}) + c(\boldsymbol{u}; \widetilde{\boldsymbol{u}}, \widehat{\boldsymbol{\varphi}}) + b(\widehat{\boldsymbol{\varphi}}, \widetilde{p}) \\
&= (\bar{\boldsymbol{l}}_1, \widehat{\boldsymbol{\varphi}}) && \forall \widehat{\boldsymbol{\varphi}} \in \boldsymbol{H}^1_0(\Omega_\eta)\,, \\
b(\widetilde{\boldsymbol{u}}, \widehat{\psi}) &= (\bar{l}_2, \widehat{\psi}) && \forall \widehat{\psi} \in L^2(\Omega_\eta)\,, \\
(\widetilde{\boldsymbol{u}}, \widehat{\boldsymbol{s}})_{\Gamma^*} &- (\mathbf{n_0}\widetilde{g}_n, \widehat{\boldsymbol{s}})_{\Gamma^*} = (\bar{\boldsymbol{l}}_3, \widehat{\boldsymbol{s}})_{\Gamma^*} && \forall \widehat{\boldsymbol{s}} \in \boldsymbol{H}^{-1/2}(\Gamma^*)\,, \\
(\widetilde{\boldsymbol{u}}, \boldsymbol{s}_1) &+ \int_{\Gamma_s} \left(\boldsymbol{V}(\widetilde{\eta}) \cdot \hat{\boldsymbol{n}} \right) \left(\chi + \frac{\partial}{\partial \boldsymbol{n}} \right) (\boldsymbol{u} \cdot \boldsymbol{s}_1)\, d\boldsymbol{x} \\
&+ (\boldsymbol{n_0} J_1(\widetilde{l} - \widetilde{\eta}), \boldsymbol{s}_1) = (\bar{\boldsymbol{l}}_6, \boldsymbol{s}_1) && \forall \boldsymbol{s}_1 \in \boldsymbol{L}^2(\Gamma_s)\,, \\
\int_{\Gamma_\eta} 2(g'_n \widetilde{g}'_n &+ g_n \widetilde{g}_n)\, ds + 2 s_2 \widetilde{s}_2 = \bar{l}_8\,,
\end{aligned}
\tag{2.64}
$$

We can assume $s_2 \neq 0$ and $\widetilde{r} = \widetilde{\eta}$. For any value of $(\bar{\boldsymbol{l}}_3, \bar{\boldsymbol{l}}_6, \bar{l}_8) \in \boldsymbol{H}^1_0(\Gamma \smallsetminus \Gamma_s) \times \boldsymbol{L}^2(\Gamma_s) \times R$ we can easily compute g_n, $\widetilde{s}_2$ and the Dirichlet boundary conditions. The problem has Dirichlet boundary conditions over all the domain and therefore we can use a classical argument. Let S be the Stokes operator, $S \cdot (\boldsymbol{u}, p)^T = (\boldsymbol{f}, 0, g_n)^T$ such that

$$
S = \begin{bmatrix} A & B^* \\ B & 0 \\ \gamma_{0n} & 0 \end{bmatrix},
$$

where γ_0 is the normal trace operator.

It is easy to show that S is an isomorphism from $\boldsymbol{H}^1(\Omega) \times L^2_0(\Omega) \to \boldsymbol{H}^{-1}(\Omega) \times L^2_0(\Omega) \times \boldsymbol{H}^{-1/2}(\Gamma)$. This property can be demonstrated using the trace theorem, the ellipticity of A and the inf-sup property [43]. We define now the operator $C : \boldsymbol{H}^1(\Omega) \times \boldsymbol{H^1}(\Omega) \times \boldsymbol{H^{-1}}(\Omega)$ such that $< C(\boldsymbol{w})\boldsymbol{u}, \boldsymbol{v} >= c(\boldsymbol{w}; \boldsymbol{u}, \boldsymbol{v})$, $\forall \boldsymbol{u}, \boldsymbol{w} \in \boldsymbol{H}^1(\Omega)$, $\forall \boldsymbol{v} \in \boldsymbol{H}^1_0(\Omega)$. The operator $C(\boldsymbol{w})\boldsymbol{u}$ is continuous in $\boldsymbol{w}$ from $\boldsymbol{H}^{1/2}(\Omega)$ into $\boldsymbol{H}^{-1}(\Omega)\, \forall \boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$. Therefore C is compact from $\boldsymbol{H}^1(\Omega)$ into $\boldsymbol{H}^{-1}(\Omega)$. As a consequence, the perturbation operator defined as

$$
P(\boldsymbol{u}, p, \boldsymbol{g}) = S + \begin{bmatrix} C(\boldsymbol{u})\boldsymbol{w} + C(\boldsymbol{w})\boldsymbol{u} \\ 0 \\ 0 \end{bmatrix},
$$

is a Fredholm operator. Thus, it has a closed range and a finite-dimensional kernel.

Now we can compute the boundary stress τ_n on Γ_n and Γ_1. In particular, on the boundary Γ_{nc} we have the variation of the control $\widetilde{\tau}_c$. With $\widetilde{\tau}_n$ on Γ_1 we compute $(\widetilde{\eta}, \widetilde{\kappa})$ by solving

$$
\begin{aligned}
(\widetilde{\tau}_n, \widehat{\varphi}_{n_0}) + (\beta J_1 \widetilde{\eta}, \widehat{\varphi}_{n_0})_{\Gamma_1} + (\gamma J_1 \nabla_{t_0} \widetilde{\eta}, \nabla_{t_0} + \widehat{\varphi}_{n_0})_{\Gamma_1} & \\
+ (\gamma_1 J_1 \nabla_{t_0} \widetilde{\kappa}, \nabla_{t_0} \widehat{\varphi}_{n_0})_{\Gamma_1} = 0 & \qquad \forall \widehat{\varphi}_{n_0} \in \boldsymbol{H}_0^1(\Gamma_1)\,, \\
(\widetilde{\kappa}, \psi_1) - (\nabla_{t_0} \widetilde{\eta}, \nabla_{t_0} \psi_1) = (\bar{l}_5, \psi_1) & \qquad \forall \psi_1 \in H_0^1(I)\,.
\end{aligned}
\tag{2.65}
$$

Finally, we consider the last equation

$$
2(\eta - \eta_0)\widetilde{\eta} + 2s_0 \widetilde{s}_0 = \bar{l}_9 \qquad \forall t \in I\,.
\tag{2.66}
$$

For a given η, it is possible to solve (2.66) for any $\bar{l}_9$. Note that if $s_0 = 0$, then $\eta = c_0$ or $\eta = d_0$. The definition of $\mathcal{R}_{ad}$ and the choice of the constants c_0 and d_0 implies the existence of solutions also when $s_0 = 0$. Given any values $l_7 \in H^{-1}(\Omega_\eta)$ we can solve the Laplacian operator for r with $r = \eta$ over Γ_s and zero elsewhere. Therefore, since all the equations of $M'(\boldsymbol{z}_0)$ have closed range, then M' has closed range in $\widehat{\boldsymbol{B}_2}$. This proves the first statement.

(2.) Since the operator $M'(\boldsymbol{u}, p, p_c, \eta, q, r, s_0)$ belongs to $\mathcal{L}(\widehat{\boldsymbol{B}_1}, \widehat{\boldsymbol{B}_2})$, its kernel is a closed subspace. Since a generic linear functional $\mathcal{F}$ on a Banach space has either $Ran(\mathcal{F}) = \{\mathbb{R}\}$, or $Ran(\mathcal{F}) = \{0\}$, and since Q' is acting on the same kernel of M', then we can assert that either $Ran(Q') = \{\mathbb{R}\}$, either $Ran(Q') = \{0\}$. We consider now the operator $Q'(\varsigma) = (\mathcal{J}(\varsigma), M'(\varsigma)), \forall \varsigma \in \widehat{\boldsymbol{B}_1}$. Then, since the range of M' is closed in $\widehat{\boldsymbol{B}_2}$ and the subspace $\mathcal{J} \cdot M'$ is closed in $\mathbb{R}$, we can assert that the range of Q' is closed in $\mathbb{R} \times \widehat{\boldsymbol{B}_2}$ (a similar approach was used in [71]).

(3.) We consider the optimal solution $(\boldsymbol{u}, p, p_c, \eta, q, r, s_0)$. Assume that $Q'(\boldsymbol{u}, p, p_c, \eta, q, r, s_0)$ is onto. By the implicit function theorem we have that $\exists (\boldsymbol{u}^*, p^*, p_c^*, \eta^*, r^*) \in \mathcal{S}_{ad}$ such that $\|\boldsymbol{u} - \boldsymbol{u}^*\|_2 + \|p - p^*\|_1 + \|p_c - p_c^*\|_{\Gamma_c} + \|\eta - \eta^*\|_{2, \Gamma_1} + \|r - r^*\|_{2, I} < \epsilon$ and $\mathcal{J}(g^*, \eta^*) < \mathcal{J}(\tau_c, \eta)$, disagreeing with the hypothesis that $\boldsymbol{z}_0$ is an optimal solution. Thus, $Q'(\boldsymbol{z}_0)$ is not onto. $\qquad \square$

Following other works in literature (see, e.g. [74]), the first-order necessary condition is now straightforward since $Q'(\boldsymbol{u}, p, p_c, \eta, q, r, s_0)$ is not onto. Then, we can finally introduce the following theorem.

Theorem 2.9. *Let $\boldsymbol{z}_0 = (\boldsymbol{u}, p, g_n, \tau_c, \eta, \kappa, r, s_0, s_2) \in \mathcal{S}_{ad}$ be a solution of the optimal control problem. Then, there exist a nonzero Lagrange multiplier*

$(\Lambda, \boldsymbol{z}_a) = (\Lambda, \boldsymbol{u}_a, p_a, \boldsymbol{g}_a, \boldsymbol{\tau}_a, \eta_a, \kappa_a, \boldsymbol{r}_a, s_{0a}, s_{2a}) \in \mathbb{R} \times \widehat{\boldsymbol{B}}_{\boldsymbol{2}}^*$, *satisfying*

$$\Lambda \mathcal{J}'(\tau_c, \eta) \cdot \widetilde{\boldsymbol{z}} + \left\langle \boldsymbol{z}_a \,, M'(\boldsymbol{z}_0) \cdot \widetilde{\boldsymbol{z}} \right\rangle = 0\,, \tag{2.67}$$

for all $\widetilde{\boldsymbol{z}} = (\widetilde{\boldsymbol{u}}, \widetilde{p}, \widetilde{g}_n, \widetilde{\tau}_c, \widetilde{\eta}, \widetilde{\kappa}, \widetilde{r}, \widetilde{s}_0, \widetilde{s}_2) \in \boldsymbol{B}_1$, *where* $< \cdot, \cdot >$ *denotes the duality pairing between* $\boldsymbol{B}_2$ *and* $\boldsymbol{B}_{\boldsymbol{2}}^*$.

With the Theorem 2.9 all the concepts for the first order necessary condition and the Lagrange multiplier approach have been reported. Now we introduce the optimality system that have to be implemented and solved in order to find the optimal solution.

2.2.5 The optimality system

As mentioned above, we now introduce some Theorems which constitute the mathematical formulation of the optimality system. The obtained equations are discretized and solved numerically with a finite element based code. Now, similarly to the Theorem 2.2 for the presented simple case, we introduce the optimality system.

Theorem 2.10. *Let* $(\boldsymbol{u}, p, g_n, \tau_c, \eta, \kappa, r, s_2, s_0) \in \boldsymbol{B_1}$ *denote a solution of the optimal control problem. When* $s_0 \neq 0$ *and* $s_2 \neq 0$, *the control* τ_c *is solution of*

$$\int_{\Gamma_c} (\lambda \tau_c + \boldsymbol{u}_a \cdot \hat{\boldsymbol{n}}) q\, dx = 0 \qquad \forall q \in L^2(\Gamma_{nc})\,. \tag{2.68}$$

Moreover $s_{0a} = 0$, $s_{2a} = 0$ *and* $(\boldsymbol{u}_a, p_a) \in H^1(\Omega) \times L^2(\Omega)$ *is solution of the adjoint problem*

$$
\begin{aligned}
\nu a(\widetilde{\boldsymbol{\varphi}}, \boldsymbol{u}_a) &+ c(\widetilde{\boldsymbol{\varphi}}; \boldsymbol{u}, \boldsymbol{u}_a) + c(\boldsymbol{u}; \widetilde{\boldsymbol{\varphi}}, \boldsymbol{u}_a) + b(\widetilde{\boldsymbol{\varphi}}, p_a) \\
&+ (\widetilde{\varphi}_{n_0}, (\eta - \eta_d))_{\Gamma_d} + (\gamma J_1 \nabla_{t_0} \widetilde{\varphi}_{n_0}, \nabla_{t_0} \eta_a)_{\Gamma_s} \\
&+ (\gamma_1 J_1 \nabla_{t_0} \widetilde{\varphi}_{n_0}, \nabla_{t_0} \kappa_a)_{\Gamma_s} + (\beta J_1\ \widetilde{\varphi}_{n_0}, \eta_a)_{\Gamma_s} = 0 \quad \forall \widetilde{\varphi} \in H^1_{\mathbf{t}_0}(\Omega_\eta)\,, \\
b(\boldsymbol{u}_a, \widetilde{\psi}) &= 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\ \ \forall \widetilde{\psi} \in L^2_0(\Omega_\eta)\,, \\
(\nabla_{t_0} \widetilde{\kappa}, \nabla_{t_0} \eta_a)_I &= (\widetilde{\kappa}, \kappa_a)_I \qquad\qquad\qquad\qquad\qquad\ \forall \widetilde{\kappa} \in H^1_0(I)\,.
\end{aligned}
\tag{2.69}
$$

If $s_0 = 0$, *the displacement* η *is determined and we have* $\eta = c_0$ *or* $\eta = d_0$. *If* $s_2 = 0$ *the stress boundary control* τ_c *is defined by a boundary velocity field* $\mathbf{g} \in \mathbf{H}^1(\Gamma_{nc})$ *that must satisfy* $\|\mathbf{g}\|_{1,\Gamma_\eta} = g_m$ *where* g_m *is the limit on the norm of the inflow boundary velocity.*

Proof. We rewrite (2.67) as

$$
\begin{aligned}
&\Lambda \left(\int_{\Gamma_d} \widetilde{\eta}(\eta - \eta_d)\, ds + \lambda \int_{\Gamma_c} \widetilde{\tau}_c\, \tau_c\, ds \right) \\
&+ \nu a(\widetilde{\boldsymbol{u}}, \boldsymbol{\omega}) + c(\widetilde{\boldsymbol{u}}; \boldsymbol{u}, \boldsymbol{\omega}) + c(\boldsymbol{u}; \widetilde{\boldsymbol{u}}, \boldsymbol{\omega}) + b(\boldsymbol{\omega}, \widetilde{p}) + (\beta J_1 \widetilde{\eta}, \omega_{n_0})_{\Gamma_s} \\
&+ (\gamma J_1 \nabla_{t_0} \widetilde{\eta}, \nabla_{t_0} \omega_{n_0})_{\Gamma_s} + (\gamma_1 J_1 \nabla_{t_0} \widetilde{\kappa}, \nabla_{t_0} \omega_{n_0})_{\Gamma_s} \\
&+ b(\widetilde{\boldsymbol{u}}, p_a) + (\widetilde{\boldsymbol{u}}, \boldsymbol{g}_a)_{\Gamma_\eta \smallsetminus \Gamma_s} - (\mathbf{n_0}\widetilde{g}_n, \boldsymbol{g}_a)_{\Gamma_\eta \smallsetminus \Gamma_s} \\
&+ (\boldsymbol{n}_0 \widetilde{\tau}_n, \boldsymbol{\tau}_a)_{\Gamma_n} + \nu a(\widetilde{\boldsymbol{u}}, \boldsymbol{\tau}_a) + c(\widetilde{\boldsymbol{u}}; \boldsymbol{u}, \boldsymbol{\tau}_a) + c(\boldsymbol{u}; \widetilde{\boldsymbol{u}}, \boldsymbol{\tau}_a) \\
&+ b(\boldsymbol{\tau}_a, \widetilde{p}) - (\widetilde{\kappa}, k_a)_I + (\nabla_{t_0} \widetilde{\eta}, \nabla_{t_0} k_a)_I + a(r, r_a) \\
&+ (\widetilde{\boldsymbol{u}}, \boldsymbol{s}_a)_{\Gamma_s} + \int_{\Gamma_s} \big(\boldsymbol{V}(\widetilde{\eta}) \cdot \hat{\boldsymbol{n}}\big)\Big(\chi + \frac{\partial}{\partial \boldsymbol{n}}\Big)\widehat{\boldsymbol{u}} \cdot \boldsymbol{s_a}\, ds \\
&+ (\boldsymbol{n}_0 J_1(\widetilde{r} - \widetilde{\eta}), \boldsymbol{s}_a)_{\Gamma_s} + \int_{\Gamma_\eta} 2 s_{2a}(g_n' \, \widetilde{g}_n' + g_n \, \widetilde{g}_n)\, ds \\
&+ 2 s_2 \widetilde{s}_2 s_{2a} + \int_I \Big(2 s_{0a}(\eta - \eta_0)\widetilde{\eta} + 2 s_0 \widetilde{s}_0 s_{0a}\Big) ds = 0\,,
\end{aligned}
\tag{2.70}
$$

By regrouping the different variations and setting to zero their contributions we have

$$
\begin{aligned}
&(\widetilde{\eta}, (\eta - \eta_d))_{\Gamma_d} + (\beta J_1 \widetilde{\eta}, \omega_{n_0})_{\Gamma_s} + (\gamma J_1 \nabla_{t_0} \widetilde{\eta}, \nabla_{t_0} \omega_{n_0})_{\Gamma_s} \\
&\quad + (\nabla_{t_0} \widetilde{\eta}, \nabla_{t_0} k_a)_I + \Big(\boldsymbol{V}(\widetilde{\eta}), \boldsymbol{n}\Big(\chi + \frac{\partial}{\partial \boldsymbol{n}}\Big)(\boldsymbol{u} \cdot \boldsymbol{s}_a)\Big)_{\Gamma_s} \\
&\quad + 2(J_1 s_{0a}(\eta - \eta_0), \widetilde{\eta})_{\Gamma_s} - (\boldsymbol{n}_0 J_1 \, \widetilde{\eta}, \boldsymbol{s}_a)_{\Gamma_s} = 0 \qquad && \forall \widetilde{\eta} \in H_0^1(I)\,, \\
&a(\widetilde{r}, r_a) - (\boldsymbol{n}_0 J_1 \widetilde{r}, \boldsymbol{s}_a)_{\Gamma_s} = 0 \qquad && \forall \widetilde{r} \in H_0(\Omega)\,, \\
&\lambda \int_{\Gamma_c} \widetilde{\tau}_c\, \tau_c\, dx + (\boldsymbol{n}_0 \widetilde{\tau}_n, \boldsymbol{\tau}_a)_{\Gamma_n} = 0 \qquad && \forall \widetilde{\tau}_c \in L^2(\Gamma_{nc})\,, \\
&\nu a(\widetilde{\boldsymbol{u}}, \boldsymbol{u}_a + \boldsymbol{\tau}_a) + c(\widetilde{\boldsymbol{u}}; \boldsymbol{u}, \boldsymbol{u}_a + \boldsymbol{\tau}_a) + c(\boldsymbol{u}; \widetilde{\boldsymbol{u}}, \boldsymbol{u}_a + \boldsymbol{\tau}_a) \\
&\quad + b(\widetilde{\boldsymbol{u}}, p_a) + (J_1 \widetilde{\boldsymbol{u}}, \boldsymbol{s}_a)_{\Gamma_s} + (\widetilde{\boldsymbol{u}}, \widehat{\boldsymbol{g}}_a)_{\Gamma_\eta \smallsetminus \Gamma_s} \qquad && \forall \widetilde{\boldsymbol{u}} \in H^1(\Omega)\,, \\
&b(\boldsymbol{\tau}_a + \boldsymbol{u}_a, \widetilde{p}) = 0 \qquad && \forall \widetilde{p} \in L_0^2(\Omega)\,, \\
&(\gamma_1 J_1 \nabla_{t_0} \widetilde{\kappa}, \nabla_{t_0} \omega_{n_0})_{\Gamma_s} - (\widetilde{\kappa}, k_a)_I = 0 \qquad && \forall \widetilde{\kappa} \in H_0^1(I)\,, \\
&-(\mathbf{n_0}\widetilde{g}_n, \widehat{\boldsymbol{g}}_a)_{\Gamma_\eta \smallsetminus \Gamma_s} + s_{2a} \int_{\Gamma_\eta} 2(g_n' \, \widetilde{g}_n' + g_n \, \widetilde{g}_n)\, ds = 0 \qquad && \forall \widetilde{g}_n \in H_0^1(\Gamma)\,.
\end{aligned}
$$

If we define $\boldsymbol{u}_a = \boldsymbol{\omega} + \boldsymbol{\tau}_a$, $\eta_a = \omega_{n_0}$ and $k_a = \gamma_1 \kappa_a$ we obtain the desired result in (2.69). We remark that in the substitution we have taken $\boldsymbol{s}_a \cdot \mathbf{t}_0 = 0$ and $\boldsymbol{g}_a \cdot \mathbf{t}_0 = 0$. The boundary derivatives term with curvature is equal to zero since the velocity vanishes on Γ_s. The term with normal derivatives is

zero since we can always extend the $\mathbf{u} \cdot \mathbf{s}_a = 0$ in the normal direction with a constant zero extension. For details see [86]. Since $s_0 \neq 0$ and $s_2 \neq 0$ the Lagrange multiplier s_{0a} and s_{2a} are identically zero and the corresponding terms vanish. □

2.3   Numerical results

In this section, we report some numerical results obtained by using the mathematical model shown in the previous sections. We implemented a standard steepest descent algorithm in the multigrid finite element code FEMuS introduced above. The numerical implementation of the system presented in the last sections is not straightforward, since the optimality system is highly nonlinear and doubles the unknowns of a standard simulation. In this work, a segregated approach is used, by splitting the solution of the state and adjoint equations to combine the result in the control gradient equation.

Algorithm 3 Description of the Steepest Descent algorithm.

---
  1. Set a state $(\boldsymbol{u}^0, p^0, \boldsymbol{\eta}^0)$ satisfying (2.26) ▷ *Setup of the state (reference case)*
  2. Compute the functional $\mathcal{J}^0$ in (2.22)
  3. Set $r^0 = 1$
  for $i = 1 \to i_{max}$ **do**
    4. Solve the system to obtain the adjoint state $(\boldsymbol{u}_a^i, p_a^i)$
    5. Set the control update $\delta p^i = -(p_c^{i-1} + \boldsymbol{u}_a^i \cdot \mathbf{n}/\lambda)$
    6. Set $r^i = r^0$
    while $\mathcal{J}^i(p_c^{i-1} + r^i \delta p^i) > \mathcal{J}^{i-1}(p_c^{i-1})$ **do**   ▷ *Line search*
      7. Set $r^i = \rho \, r^i$
      8. Solve (2.26) for the state $(\boldsymbol{u}^i, p^i, \boldsymbol{\eta}^i)$ with $p_c^i = p_c^{i-1} + r^i \delta p^i$
      if $r^i < toll$ **then**
        Line search not successful   ▷ *End of the algorithm*
      end if
    end while
  end for

In Algorithm 3 a detailed description of the Steepest Descent method used for the iterative solution of the optimality system. After an initial setup, where the initial state 1. and the initial functional value 2. are computed, the algorithm consists of two nested loops. In the outer loop, the adjoint system is solved 4. together with the control equation 5. to obtain the gradient

direction δp. In the inner loop, a backtracking line search is used (see [75]). The algorithm stops either when the step length r^i is lower than a minimum value *toll*, either when two functional computed consecutively are similar. This means that no more improvements can be obtained. In this work we consider *toll* $= 10^{-8}$. Note that, for simplicity, in this section we consider the pressure p_c instead of the controlled variable τ_c as introduced in the last sections. Since we set to zero the inlet/outlet tangential velocity, in two-dimensional domains the term independent of the pressure in the normal stress $\boldsymbol{\tau}_s$ vanishes. Therefore, τ_c and p_c are mutually dependent.

Different numerical tests are reported in the following sections to show the robustness of all the developed algorithms. In particular, a simple case where Ω_d reduces to only one point is reported in order to evaluate the effectiveness of the control algorithm in the simplest case. Then, some cases on the grid convergence of the method are shown.

2.3.1 Zero-dimensional desired domain

As mentioned above, the first test presented in this section is the simplified case where the domain Ω_d, in which the variable η is required to be similar to the desired one η_d, is reduced to a single point. We consider a rectangular domain $\Omega = \{(x, y) : x \in [0, 0.1], y \in [0, 0.3]\}$ as shown in Figure 2.3.
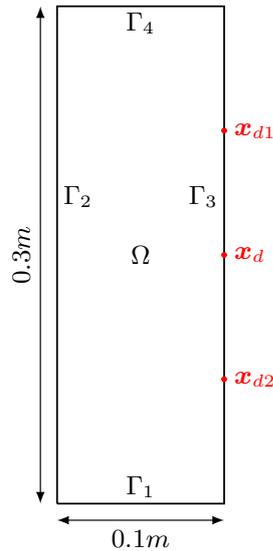
Figure 2.3: Geometry and controlled points ($\boldsymbol{x}_d$, $\boldsymbol{x}_{d1}$ and $\boldsymbol{x}_{d2}$) of the cases tested in this section.

Displacement reduction

In the first studied case, the fluid flows vertically from the bottom to the top, the region of the boundary Γ_2 represents a solid wall with no-slip boundary condition ($\boldsymbol{u} = \boldsymbol{0}$) and Γ_3 is the membrane where the generalized Robin boundary condition is imposed. The fluid has density $\rho^f = 1000 \ kg/m^3$ and dynamic viscosity $\mu = 100 \ Pa \cdot s$, and for the approximation of the solid to mono-dimensional membrane we consider $\beta = 60kPa/m$ and thickness $h_s = 0.0075 \ m$. For the simulations, the domain was uniformly divided with a regular rectangular mesh.

The simulations aim to control the displacement of the point $\boldsymbol{x}_d$ of the membrane optimizing the pressure of the fluid on Γ_i. Note that the desired displacement field is directed along the x-axis, following the prescriptions of the Koiter model introduced in the previous sections. We consider first the uncontrolled case (equivalent to $\lambda \to \infty$) with a prescribed inlet pressure of $6000Pa$. Under this hypothesis, the controlled point shows a displacement of $\eta = 0.015824 \ m$. We impose $\eta_d = 0.005 \ m$, therefore the control algorithm should act to reduce η to the value η_d, changing the pressure of the fluid on Γ_i. The objective functional of the problem reads

$$\mathcal{J}(\eta, p_c) = \frac{1}{2}(\eta - \eta_d)^2 + \frac{\lambda}{2}\int_{\Gamma_i} p_c^2 \, d\Gamma \,. \tag{2.71}$$

We first study the behavior of the functional $\mathcal{J}$ depending on the value of the regularization parameter λ. As can be noted in Table 2.1, the smaller is λ, the closer the displacement of the controlled point $\boldsymbol{x}_d$ is to the desired one. This result is expected, since with larger λ the contribution of the regularization term in the minimization of the functional is more relevant. Therefore, with larger λ we find more regular optimization parameter p, but less precise displacement η. In general, all the control problem simulated show an improvement of the initial objective functional $\mathcal{J}$, thus a solution closer to the objective is found, as can be see also from the reported η values on $\boldsymbol{x}_d$ (reported as η_{opt} in Table). At the same time, the lower is the λ value, the higher is the number of iterations of the optimality algorithm necessary to find the optimal solution, as reported in Figure 2.4. This results in an increase of the computational cost of the whole algorithm. Note that the solution converges to the requested solution η_d with the iterations of the algorithm. Note also that the only case in which the solution doesn't converge to η_d is with $\lambda = 10^{-8}$. In fact, in such case the algorithm is not able to find

Table 2.1: Objective functional $\mathcal{J}$, displacement η and number of iteration obtained with different λ values.

λ	$\mathcal{J}(\eta, p_c)$	$\eta_{opt}[m]$	Iterations
∞	$5.85839 \cdot 10^{-05}$	0.015824	–
10^{-08}	$2.19246 \cdot 10^{-06}$	0.002906	4
10^{-09}	$5.54438 \cdot 10^{-09}$	0.004895	8
10^{-10}	$2.18941 \cdot 10^{-10}$	0.004979	10
10^{-11}	$6.10506 \cdot 10^{-12}$	0.004997	12
10^{-12}	$3.86734 \cdot 10^{-15}$	0.005000	26
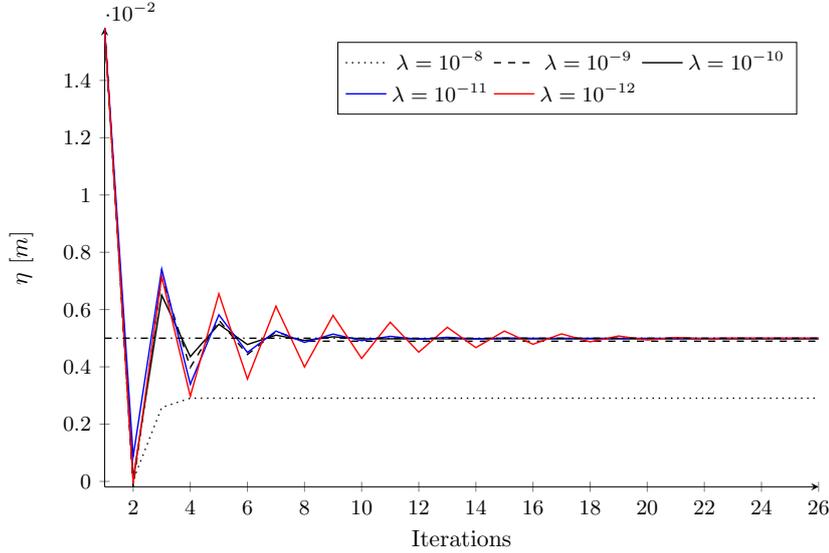
Figure 2.4: Value of $\eta = dx$ on the point $\boldsymbol{x}_d$ depending on the algorithm iteration number, for different value of the regularization parameter.

a better solution than $\eta = 0.002906m$ able to reduce the functional, since the regularization term affects strongly the functional value.

We focus now on the controlled inlet pressure field. In fact, depending on the regularization parameter different inlet pressure fields are obtained. In Figure 2.5 the controlled pressure field along the boundary Γ_1 is reported. Note that the choice of the regularization parameter strongly affects the controlled pressure field. With less regularization, the objective term dominates in the functional and the pressure can have larger values, thus effectively controlling the membrane displacement. In Figure it is also reported the ref-
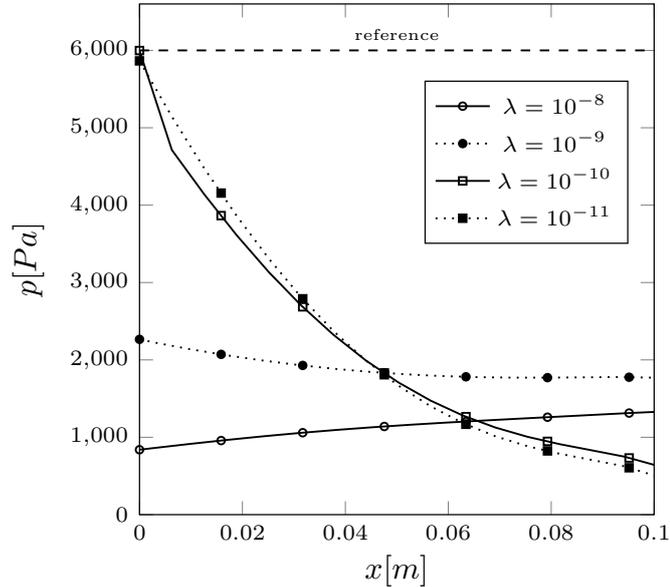
Figure 2.5: Control pressure p on Γ_1 with different regularization parameters. The dotted line represents the pressure in the reference case with no control (i.e. $\lambda = \infty$).

erence starting pressure. The comparison between the uncontrolled field and all the controlled pressure fields shows that the control is strongly modifying the solution on $\Gamma_c = \Gamma_1$ in order to obtain the desired displacement η_d.

In conclusion, the presented results show that the algorithm strongly affects the uncontrolled simulation in order to find the optimal solution. While the pressure is directly modified by the algorithm, all the other variables (e.g. the velocity) are different from the reference case. It is important to tune properly the regularization term: high values of λ lead to incorrect solutions, low values of it lead to higher computational costs.

Similar numerical tests have been developed with a displacement increase (instead of the reduction presented in this paragraph). The results are similar to the one presented in this section and, for this reason, are not shown in this work.

Airbag with the desired region on two points

We consider now the same geometry presented in Figure 2.3, with the same physical properties. We consider airbag-like boundary conditions, with a no-slip condition imposed on Γ_1 and Γ_4 and the Koiter boundary condition on

Γ_3. The controlled pressure is imposed on $\Gamma_c = \Gamma_2$. In this framework we want to control the displacement field on two different points, $\boldsymbol{x}_{d1}$ and $\boldsymbol{x}_{d2}$ in Figure. This test is carried out in order to show the advantage to use the optimal control approach, since with standard control techniques it is difficult to achieve non-constant desired displacements, since the prescribed pressure on Γ_c is often constant. In particular, with a *try and fail* approach, the solution of the current problem could lead to high computational costs. We impose that $\eta_{d1} = -0.01m$ and $\eta_{d2} = 0.01m$, therefore we are simultaneously imposing an increase and a reduction of the displacement in the reference case on two different points.
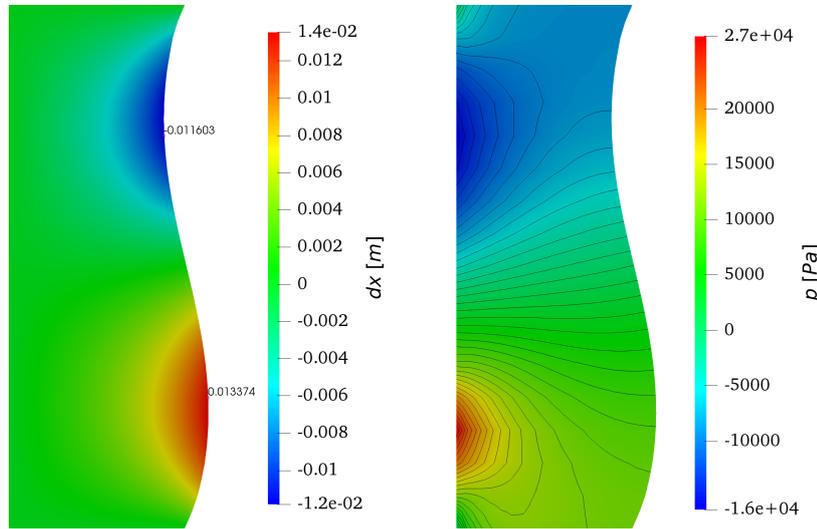
Figure 2.6: Qualitative behavior of the displacement dx (left) and the pressure p (right) in Ω with control on $\boldsymbol{x}_{d1}$ and $\boldsymbol{x}_{d2}$.

In Figure 2.6 we report the displacement and the pressure fields for the presented case for $\lambda = 10^{-10}$. Note that the two different requirements on the two points implies the imposition of a negative pressure on Γ_2. In Figure 2.6 the displacement field doesn't match perfectly the requirements of $\eta_{d1} = -0.01m$ and $\eta_{d2} = 0.01m$, however the general behavior of the membrane seems to reproduce the required profile of the surface. The same concept can be seen in Table 2.2, where the values of η_{d1} and η_{d2} doesn't match the requirements. However, the values of the functional (e.g. the distance from the objective) reduces with respect to the non-controlled case ($\lambda = \infty$), therefore an improvement is achieved.

In Figure 2.7 the comparison between the imposed pressure fields on Γ_2

Table 2.2: Objective functional $\mathcal{J}$, displacement on the two controlled points and number of iteration obtained with different λ values.

λ	$\mathcal{J}(\eta, p_c)$	$\eta_{1opt}[m]$	$\eta_{2opt}[m]$	Iterations
∞	$1.01327 \cdot 10^{-4}$	0.001152	0.001152	$-$
10^{-8}	$2.21289 \cdot 10^{-5}$	-0.010695	0.005348	4
10^{-10}	$1.39504 \cdot 10^{-5}$	-0.011603	0.013374	6
10^{-12}	$2.54909 \cdot 10^{-5}$	-0.010032	0.004951	4
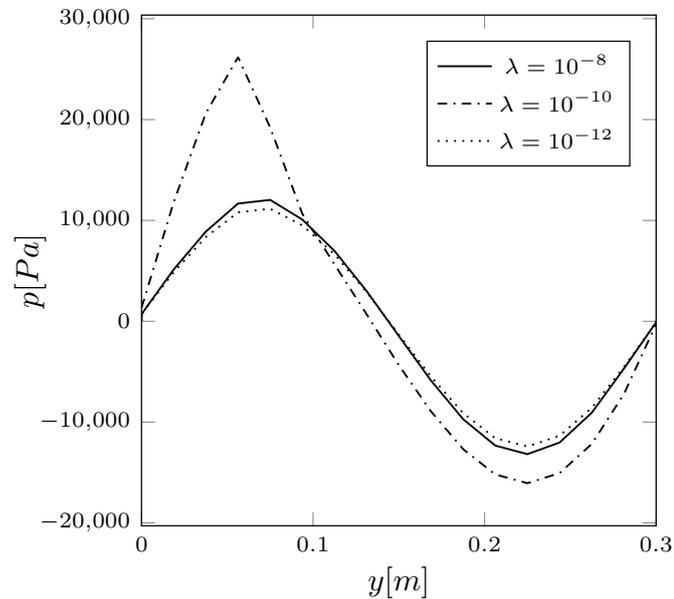
Figure 2.7: Controlled pressure field on Γ_2 with the objective $\eta_{d_1} = -0.01 \; m$ and $\eta_{d_2} = 0.01 \; m$ for different values of λ.

is reported. Again, note that the value of λ affects the pressure field to be imposed on the control domain Γ_c. We remark that higher values of λ lead to more regular solutions. In general, this case is an ill-posed control problem, due to various local minima close to each other.

The dependence of the optimal solution on λ is an interesting topic, and will be further studied in future works. In this work we will not further analyze the order of magnitude of λ, since the algorithm finds always an improved solution, for all the tested values of λ.

2.3.2   Grid convergence

As introduced in the last chapter, we show now some results on the grid convergence of the proposed method. In particular, we require the reduction of the integral of the difference between η and η_d on the domain where the control acts Ω_d with the grid refinement. We report now different tests, depending on the requested displacement and on the size of Ω_d.
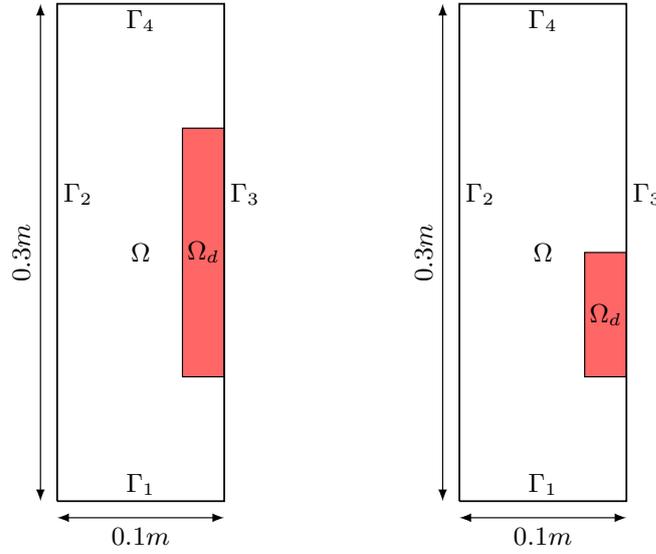
Figure 2.8: The two domains used for the grid convergence tests: with Ω_d of dimension $0.025m \times 0.15m$ (left) and $0.025m \times 0.075m$ (right).

In all the considered cases we use the same physical values: fluid density $\rho^f = 1000\ kg/m^3$, fluid dynamic viscosity $\mu = 100\ Pa \cdot s$, and for the approximation of the solid to mono-dimensional membrane we consider $\beta = 60kPa/m$ and thickness $h_s = 0.0075\ m$. Moreover, we will consider $\lambda = 10^{-10}$, unless stated otherwise. We consider a 2×2 mesh of the domains introduced in Figure 2.8. The control domain Γ_c is equivalent to the surface Γ_1, and we refine it with a multi-grid approach. Interested reader in such a technique can see [65, 57] as reported also in section 1.5.3.

Displacement reduction

We first compute a displacement reduction test for the plane channel in Figure 2.8 on the left. In particular, we request a desired displacement of $\eta_d = 0.005m$ on Ω_d. We control the pressure p field over Γ_1. Note that the average

displacement on Ω_d in the uncontrolled case is $\bar{\eta}_d = 0.00751m$, then we are simulating a displacement reduction control.
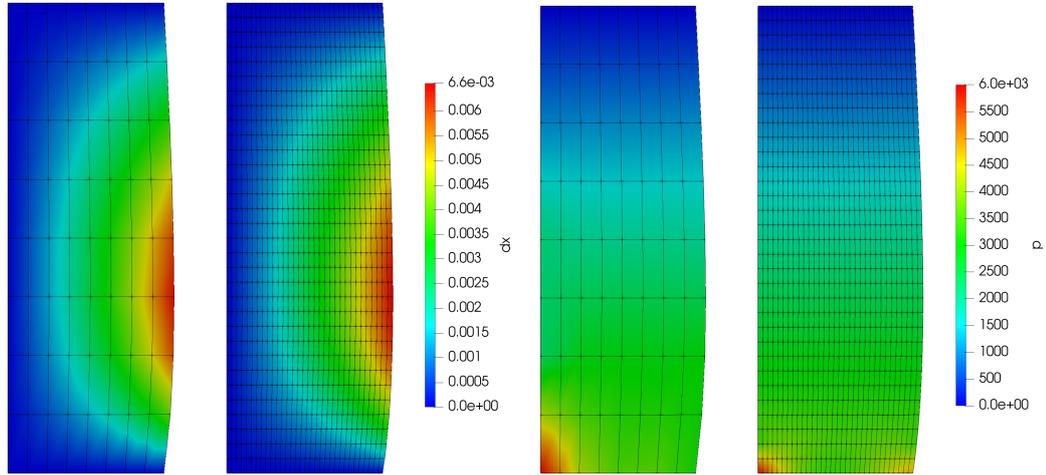
Figure 2.9: Comparison between the control with 2 and 4 levels of refinement. We report the displacement fields dx (right) and the pressure fields (left).

In Figure 2.9 on the left, the controlled displacement fields for the considered mesh refined 2 and 4 times is reported. Note that the grid reported in Figure reports the quadratic elements as consisting of four elements. For example, the 2-levels mesh is composed of a 4×4 grid. Also the pressure fields are reported on the right in the Figure. It can be noted that the grid refinement affects the solution of the numerical problem and all the involved fields.

In Figure 2.10 the comparison between the controlled pressure field is reported along the controlled boundary Γ_1. In particular, all the presented solutions seem to converge with the grid to a certain pressure field. All the fields differ from the reference pressure field, and in particular it can be noted that the pressure is reduced by the control algorithm at the inlet. This result is expected since we are studying a reduction of the displacement on Ω_d.

In Figure 2.11 the comparison between the velocity fields over Γ_1 is reported. Note that the control pressure influences all the variables. In particular, the velocity field is reduced with respect to the reference velocity, due to the lower imposed pressure.

In Table 2.3 the values of the distance between the desired solution and the solution find with the optimal control algorithm is reported. In particular, we compute the objective distance as $\int_{\Omega_d} (\eta - \eta_d)^2 \, dx$, where the integral over
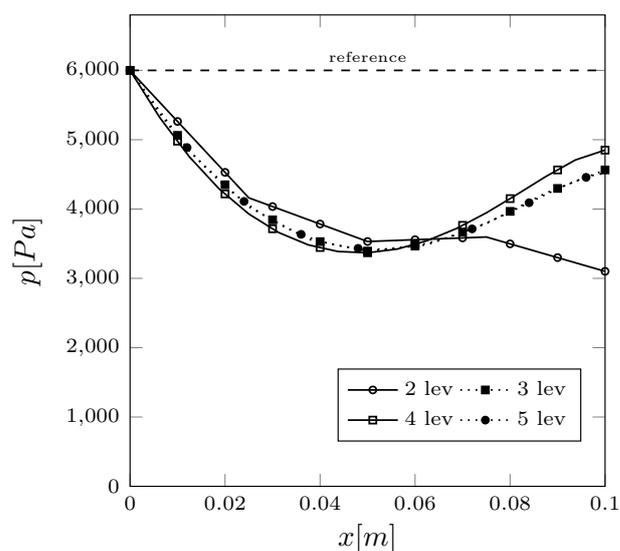
Figure 2.10: Comparison between the pressure field over $\Gamma_c = \Gamma_1$ for different mesh refinements.
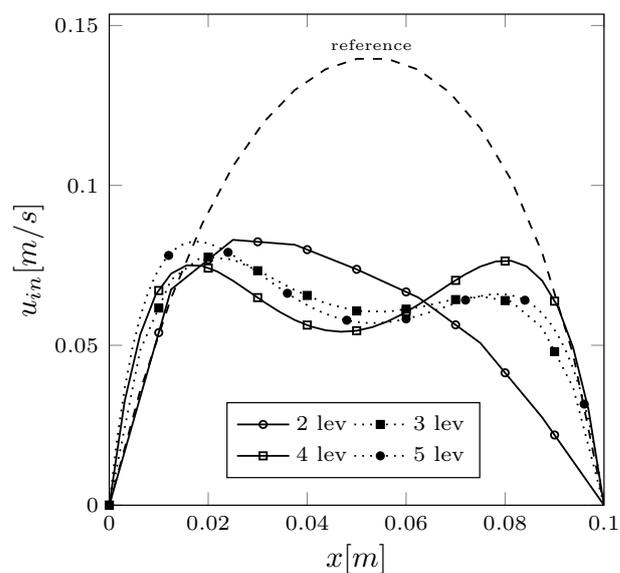
Figure 2.11: Comparison between the velocity field over $\Gamma_c = \Gamma_1$ for different mesh refinements.

Ω_d has been developed on the same grid with the same quadrature rule. Note that the distance from the objective decrease with the refinement of the grid. This is a good indicator of the grid convergence of the algorithm since

Table 2.3: Distance from the objective η_d depending on the number of refinement levels.

Levels	λ	$\int_{\Omega_d}(\eta - \eta_d)^2\, dx$	Iterations	R
2	∞	$3.45 \cdot 10^{-8}$	–	–
2	10^{-10}	$4.01 \cdot 10^{-9}$	10	$1.16 \cdot 10^{-1}$
3	10^{-10}	$3.41 \cdot 10^{-9}$	10	$9.88 \cdot 10^{-2}$
4	10^{-10}	$3.26 \cdot 10^{-9}$	12	$9.45 \cdot 10^{-2}$
5	10^{-10}	$3.22 \cdot 10^{-9}$	10	$9.32 \cdot 10^{-2}$

refined solutions have a better result. In Table, we also report the number of iterations of the algorithm to find the optimal solution. All these value are compared with the reference uncontrolled simulation ($\lambda = \infty$). We also report the reduction rate, defined as

$$R = \frac{\int_{\Omega_d}(\eta - \eta_d)^2\, dx}{\int_{\Omega_d}(\bar{\eta}_d - \eta_d)^2\, dx}\,. \tag{2.72}$$

In Table, the values of R are always $R < 1$, therefore the solution is always improved with respect to the reference one. This means that the control algorithm finds always a solution better then the reference one.

Displacement increase

Now a test on a displacement increase with respect to the reference configuration is introduced. In particular, we request a desired displacement of $\eta_d = 0.02m$ on Ω_d. We consider the Ω_d domain highlighted in Figure 2.8 on the left. All the physical quantities are considered equal to the previous case. The average displacement field in Ω_d in the uncontrolled case is $\bar{\eta}_d = 0.00833m$.

In Figure 2.12 the displacement field over the objective region Ω_d is reported. Note that the control acts in order to increase the displacement field to the desired values. Note that the values of $\eta = dx$ on the studied domain are around the desired value of $\eta_d = 0.02m$. In Figure 2.13 the adjoint velocity field is reported for number of levels 4 (on the left) and 5 (right). The adjoint velocity is obtained through the solution of the adjoint system, and it is used to solve the control equation. Note that the adjoint system leads to a
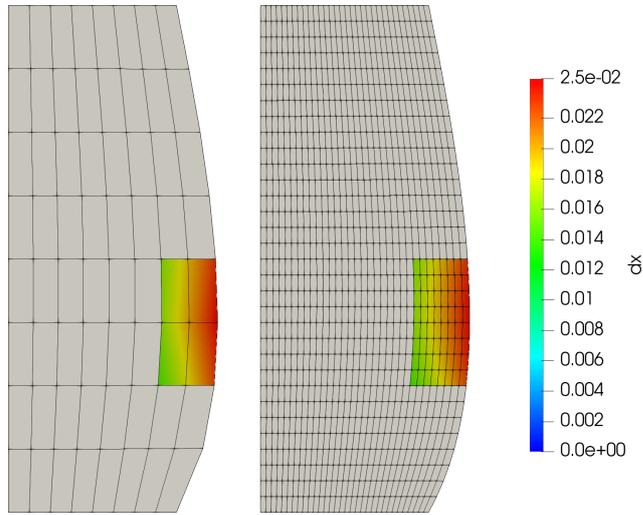
Figure 2.12: Displacement field on Ω_d for Lev = 2 (left) and Lev = 4 (right).
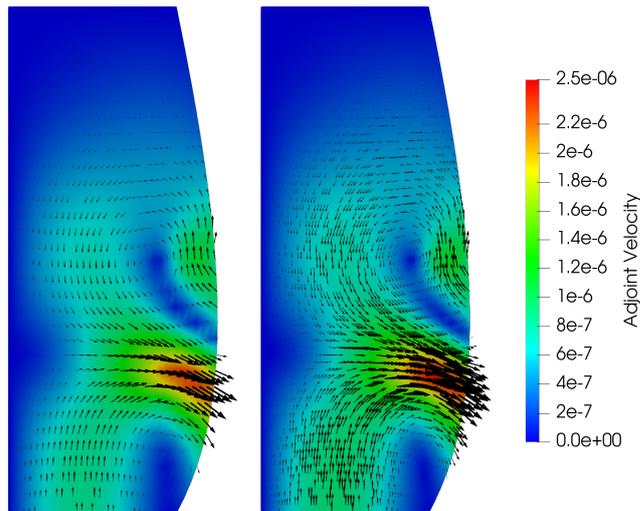
Figure 2.13: Adjoint velocity fields for Lev = 4 (left) and Lev = 5 (right).

solution similar to a classical Navier-Stokes system with consistent boundary conditions. The fluid enters from the same inlet of the state problem, and exit from the domain Ω_d. The controlled pressure p_c depends on the value of the adjoint velocity in the inlet, as reported in the previous section. Note that the adjoint velocity field seems to be consistent with the grid refinement.

In Figure 2.14 we report the comparison between the control pressure fields (over Γ_1) for different grid refinement. In particular, in contrast to the previously reported test, a meaningful increase of the controlled pressure
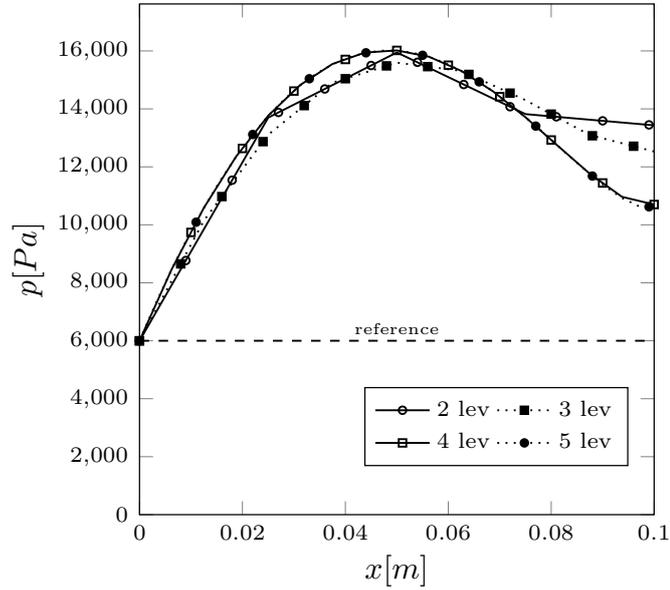
Figure 2.14: Comparison between the pressure field over $\Gamma_c = \Gamma_1$ for different mesh refinements.

can be noted in this case. Obviously, this is a direct consequence of the request for an increase of the displacement on Ω_d. Note that the control pressure converges with the grid, since the pressure fields for 4 and 5 levels are coincident.

Table 2.4: Distance from the objective η_d depending on the number of refinement levels.

Levels	λ	$\int_{\Omega_d}(\eta - \eta_d)^2\,dx$	Iterations	R
2	∞	$2.91 \cdot 10^{-7}$	—	—
2	10^{-10}	$2.39 \cdot 10^{-8}$	12	$8.21 \cdot 10^{-2}$
3	10^{-10}	$1.76 \cdot 10^{-8}$	7	$6.05 \cdot 10^{-2}$
4	10^{-10}	$1.58 \cdot 10^{-8}$	12	$5.43 \cdot 10^{-2}$
5	10^{-10}	$1.53 \cdot 10^{-8}$	11	$5.26 \cdot 10^{-2}$

In Table 2.4 the comparison between all the tested cases is reported. In particular, all the cases are compared with the reference case without control (i.e. $\lambda = \infty$). As in the previous case, we report the value of the distance from the objective calculated with the integral $\int_{\Omega_d}(\eta - \eta_d)^2\,dx$. All

the controlled simulations show a minor distance from the objective compared to the reference case. The reduction rate, introduced in (2.72), is $R < 0.1$ for all the tested grids. This means that the solution is always improved compared to the reference one, and the solutions for higher refinement levels reduces the distance from the objective. In Table we report also the number of iterations necessary to find the optimal solution for each grid refinement.

2.3.3 Variable desired field

We consider now a variable desired displacement field. This test can be useful for practical applications where a non-constant objective is required. Moreover, optimal control is a good approach to such problems, since the try and fail approaches often can't find a good solution to this optimization problem. We present two different problems, with a sinusoidal and a step desired field. We consider the geometry in Figure 2.8 (left) with $\Omega_d = \{(x, y) : x \in [0.075, 0.1], y \in [0.075, 0.225]\}$.

We consider a fluid density $\rho^f = 1000 \ kg/m^3$, fluid dynamic viscosity $\mu = 1 \ Pa \cdot s$, and for the approximation of the solid to mono-dimensional membrane we consider $\beta = 60 kPa/m$ and thickness $h_s = 0.0075 \ m$. Moreover, we will consider $\lambda = 10^{-6}$, unless stated otherwise. We consider a 2×2 mesh of the domain, refined $N_{lev} = 3$ times with a multigrid technique. The control domain Γ_c is equivalent to the surface Γ_2, so we simulate an airbag like case. We also impose an inlet pressure of $p_{in} = 600 Pa$.

A simple sinusoidal desired displacement

We start considering a simple sinusoidal case, where η_d is defined as

$$\eta_d = 0.01 + 0.0025 \sin\left(\frac{2\pi(y - 0.075)}{0.3}\right) \quad [m].$$

The shape of this function has been chosen in order to be easily reproduced using the control strategy presented in the last sections. In Table 2.5 we report the distance from the objective ε defined as

$$\varepsilon = \int_{\Omega_d} (\eta - \eta_d)^2 \, d\boldsymbol{x},$$

and the reduction rate, defined in (2.72). We also report the value of λ, that is equal to 10^{-6} in all the studied cases, with the only exception of the

Table 2.5: Distance from the objective and reduction rate R of the objective functional depending on the iteration of the control algorithm.

It	λ	ε	R
–	∞	$4.15 \cdot 10^{-7}$	–
1	10^{-6}	$1.90 \cdot 10^{-7}$	$4.59 \cdot 10^{-1}$
3	10^{-6}	$9.90 \cdot 10^{-8}$	$2.39 \cdot 10^{-1}$
5	10^{-6}	$1.37 \cdot 10^{-8}$	$3.30 \cdot 10^{-2}$
7	10^{-6}	$9.45 \cdot 10^{-9}$	$2.28 \cdot 10^{-2}$
9	10^{-6}	$7.46 \cdot 10^{-9}$	$1.80 \cdot 10^{-2}$

reference case ($\lambda = \infty$). For simplicity, we report only the odd iterations. Note that the distance from the objective decreases in all the iteration of the optimization algorithm, and reach a final value of $7.46 \cdot 10^{-9}$. The final reduction rate of $1.80 \cdot 10^{-2}$ suggests that the final solution is 55 times closer to the objective with respect to the initial solution.
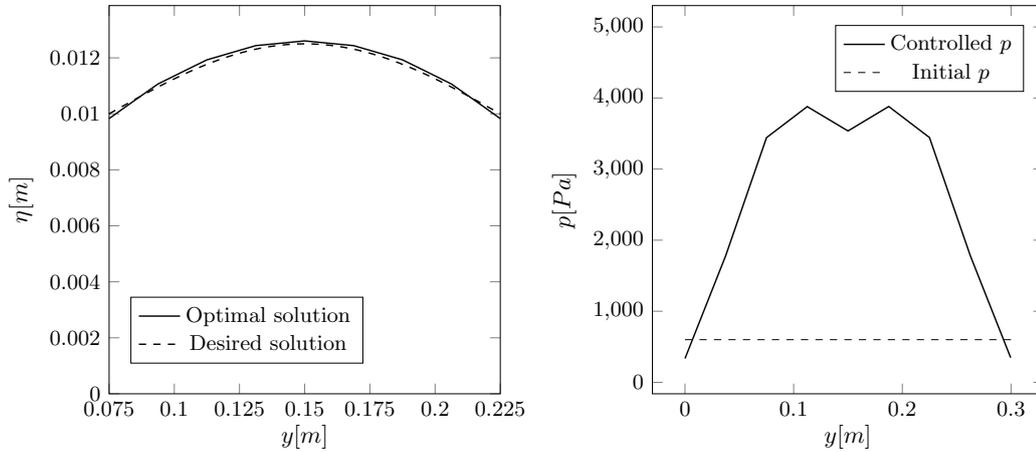
Figure 2.15: Displacement field (left) along the line between the points $(0.1, 0.075)$ and $(0.1, 0.225)$, and controlled pressure field (right) along the boundary Γ_2.

In Figure 2.15 on the left, the displacement field along the line ℓ between the points $(0.1, 0.075)$ and $(0.1, 0.225)$ (such that $\ell \in \Omega_d$) is reported. The optimal and reference solutions have a similar behavior, therefore we can

conclude that the optimization algorithm found a good solution, in perfect agreement with the desired displacement. In Figure on the right, we report the controlled pressure field along Γ_2 to have the optimal displacement reported.

Step function desired displacement

We consider now a desired step function η_d, defined as

$$
\eta_d = \begin{cases} 0.0075\, m & \text{if } y \in [0.075, 0.15[\\ 0.015\, m & \text{if } y \in [0.15, 0.225]\,. \end{cases}
$$

All the physical parameters are the same used in the last cases. The presented case is not a straightforward control problem, since it is not easy to represent a step function displacement by controlling the pressure on Γ_2. We split the distance from the objective in two terms ε_1 and ε_2, such that

$$
\varepsilon_i = \int_{\Omega_{d,i}} (\eta - \eta_d)^2 \, d\boldsymbol{x}\,, \quad i = 1, 2\,,
$$

where $\Omega_{d,1} = \{(x,y) : x \in [0.075, 0.1], y \in [0.075, 0.15[\}$ and $\Omega_{d,2} = \{(x,y) : x \in [0.075, 0.15], y \in [0.075, 0.225]\}$. Obviously, $\varepsilon = \varepsilon_1 + \varepsilon_2$.

Table 2.6: Distance from the objective for both the domains $\Omega_{d,1}$ and $\Omega_{d,2}$, and overall distance from the objective ε depending on the iteration of the algorithm. The reduction rate R of the objective functional is also reported.

It	λ	ε_1	ε_2	ε	R
$-$	∞	$3.64 \cdot 10^{-7}$	$7.63 \cdot 10^{-8}$	$4.40 \cdot 10^{-7}$	$-$
1	10^{-6}	$2.23 \cdot 10^{-8}$	$2.17 \cdot 10^{-7}$	$2.39 \cdot 10^{-7}$	$5.43 \cdot 10^{-1}$
3	10^{-6}	$1.05 \cdot 10^{-8}$	$7.97 \cdot 10^{-8}$	$9.02 \cdot 10^{-8}$	$2.05 \cdot 10^{-1}$
5	10^{-6}	$7.74 \cdot 10^{-8}$	$5.49 \cdot 10^{-9}$	$8.28 \cdot 10^{-8}$	$1.88 \cdot 10^{-1}$
7	10^{-6}	$2.31 \cdot 10^{-8}$	$1.33 \cdot 10^{-8}$	$3.46 \cdot 10^{-8}$	$7.87 \cdot 10^{-2}$
9	10^{-6}	$1.28 \cdot 10^{-8}$	$1.49 \cdot 10^{-8}$	$2.77 \cdot 10^{-8}$	$6.29 \cdot 10^{-2}$

'

In Table 2.6, ε_1 and ε_2 are reported. Note that the values of ε_1 and ε_2 are not uniformly decreasing with the algorithm iterations, but their sum does.

In fact, the algorithm is designed to reduce the value of ε: such value and, consequently, the reduction rate R, decreases to a value of $2.77 \cdot 10^{-8}$ and $6.29 \cdot 10^{-2}$ respectively after 9 iterations. The overall reduction rate is greater than the previous case, since the objective is more difficult to be reached.
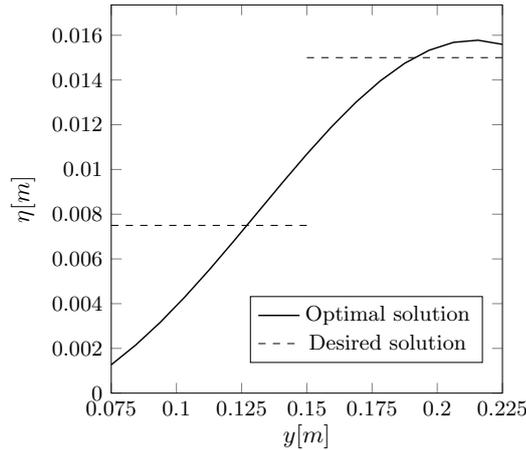
Figure 2.16: Step function desired field test for Levels= 3. Displacement field along the line between the points $(0.1, 0.075)$ and $(0.1, 0.225)$.

In Figure 2.16, the displacement field along ℓ is reported. It is very difficult to have a perfect match between the optimal solution and the objective step function η_d. The optimization algorithm works to minimize the distance between the optimal and the desired solutions, and the solution found after 9 iterations of the optimization algorithm is reported in Figure. We remark that, even if the proposed optimal solution does not represent a perfect matching with the desired one, it is an upgrade of the initial solution by $1/R \approx 16$ times.

Sinusoidal desired displacement

We finally consider a sinusoidal case difficult to be represented, defined as

$$\eta_d = 0.015 + 0.0075 \sin\left(\frac{2\pi(y - 0.075)}{0.15}\right) \quad [m].$$

Note that, in contrast to the simple sinusoidal case presented previously, in this case we request a full period of 2π, with a change in the second derivatives sign. As in the case of the step function η_d, a perfect representation of the desired displacement is not an easy task under these hypotheses.

Table 2.7: Distance from the objective and reduction rate R of the objective functional depending on the iteration of the control algorithm.

It	λ	ε	R
$-$	∞	$8.36 \cdot 10^{-7}$	$-$
2	10^{-6}	$4.66 \cdot 10^{-7}$	$5.59 \cdot 10^{-1}$
4	10^{-6}	$2.14 \cdot 10^{-7}$	$2.56 \cdot 10^{-1}$
6	10^{-6}	$1.30 \cdot 10^{-7}$	$1.55 \cdot 10^{-1}$
8	10^{-6}	$9.87 \cdot 10^{-8}$	$1.18 \cdot 10^{-1}$
10	10^{-6}	$7.09 \cdot 10^{-8}$	$8.48 \cdot 10^{-2}$

In Table 2.7 we report the values of λ, the distance from the objective ε and the reduction rate R. We report only the even iterations for simplicity. Note that the distance from the objective is reduced more than 10 times in 10 iterations of the optimization algorithm. In particular, the reduction of R is equal to $1/R \approx 12$ times, similarly to the last test for desired step function displacement.
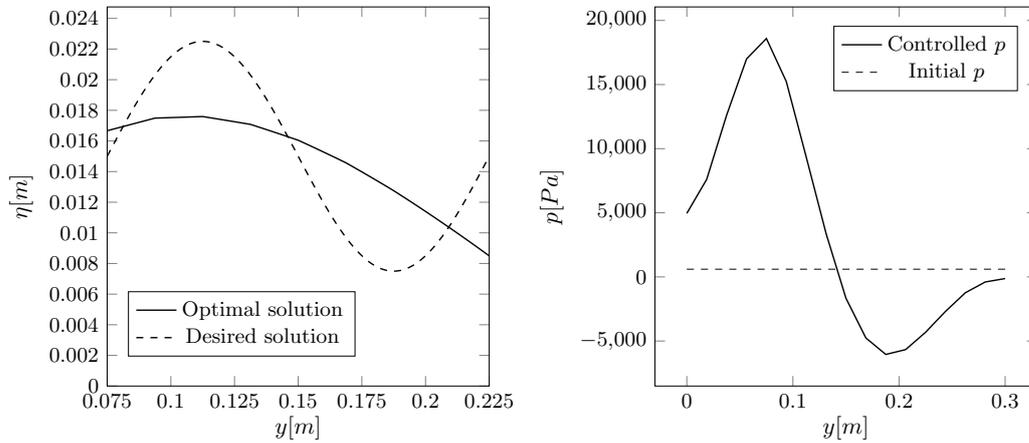
Figure 2.17: Displacement field (left) along the line ℓ, and controlled pressure field (right) along the boundary Γ_2 in the sinusoidal case.

In Figure 2.17, the displacement field along ℓ and the pressure field along Γ_2 are reported. Note that, as in the step case, the optimal solution does not match the desired one. However, the algorithm finds a solution closer to the

objective in comparison to the reference one, as can be seen also in Table 2.7. This is also due to the numerical problem: the choice of the control domain, the control variable and the domain where the desired solution is requested strongly affects the well-posedness of the numerical algorithm.

The optimal pressure field found by the algorithm is bounded. However, the regularity of the p_c found with this approach is not strong. The pressure field suggested by the optimization algorithm is difficult to be represented in real applications. The regularity can be improved by using a different approach for the regularization term. In particular, in the case of $p_c \in H^{1/2}$, the correct norm for the regularization term is $\| \cdot \|_{H^{1/2}}$. In the next chapter we introduce the fractional Sobolev spaces in order to implement such a regularization term.

Fractional models

In the adjoint optimal control theory, as it has been previously shown in Chapter 2, a regularization term is added to the cost functional $\mathcal{J}$ in order to force better mathematical and numerical properties to the optimization process. This term usually contains the L^2-norm of the control q, penalized with a parameter λ.

Boundary optimal control problems are one of the classes of greatest interest. In fact, the possibility of controlling the behavior of a physical system may often take place by changing the quantities of certain variables at the boundary of the domain. In such cases, normally one may wish to impose further requirements on the regularity of the controls, i.e. H^1 norm, to be added to the cost functional. This technique works properly in many boundary control problems, but it is not based on a coherent mathematical formulation.

Theoretical results in mathematical analysis imply that the connection between functions defined on the domain of a PDE and their restriction to the boundary gives rise to *fractional order Sobolev spaces*. The presence of these spaces induces the necessity to deal with the numerical discretization of fractional derivatives. In particular, it can be demonstrated that the natural space where optimal boundary control should be sought is $H^{1/2}(\Gamma_c)$. Note that $H^1(\Gamma_c)$ is a subspace of $H^{1/2}(\Gamma_c)$.

Alternative approaches, based on the *lifting functions* [87], can be used to not handle the fractional norms, but they will not be considered in this work.

In the following paragraphs, the fractional Sobolev spaces will be introduced. This will lead to the analysis of some numerical methods for nonlocal and fractional models.

Besides the context of boundary control problems, the attention given in recent years to fractional derivatives and their discretization has been growing. Fractional derivatives in fact appear in the mathematical modeling of forward problems whenever non-local effects are taken into account. These derivatives may involve both the time and space variables. Examples of mathematical models involving fractional derivatives are met in several fields, such as finance, fluid, and solid mechanics, stochastic modeling, quantum mechanics, and the social sciences.

All the models and the simulations presented in this chapter has been developed in a collaboration with the Department of Mathematics and Statistics of the Texas Tech University.

3.1   The Riesz fractional Laplacian

In this section the fractional Sobolev spaces and their properties are introduced. Much of this introductory section is derived from the hitchhiker's guide [13], a recommended reading for anyone who wants to start working with the fractional spaces.

It is also reported a mathematical introduction to the fractional Laplacian and his properties. For this purpose, the nonlocal models will be introduced since the fractional Laplacian can be seen as a special case of a nonlocal operator. It will be also introduced the weak formulation of such an operator, which will lead, in the following sections, to the numerical resolution of the fractional Laplacian.

3.1.1   Introduction to the fractional Sobolev spaces

We introduce now the general definition of fractional Sobolev space.

Definition 3.1. *Given Ω, a general, possibly non-smooth open set in $\mathbb{R}^n$, the fractional Sobolev space $W^{s,p}(\Omega)$ is defined as*

$$W^{s,p}(\Omega) = \left\{ u \in L^p(\Omega) : \frac{|u(\boldsymbol{x}) - u(\boldsymbol{y})|}{|\boldsymbol{x} - \boldsymbol{y}|^{\frac{n}{p}+s}} \in L^2(\Omega \times \Omega) \right\}, \qquad (3.1)$$

for any $p \in [1, +\infty)$ and with a fractional exponent $s \in (0,1)$.

Definition 3.2. *The natural norm in $W^{s,p}(\Omega)$ is defined as*

$$\|u\|_{W^{s,p}(\Omega)} = \left(\int_{\Omega} |u|^p dx + \int_{\Omega} \int_{\Omega} \frac{|u(\boldsymbol{x}) - u(\boldsymbol{y})|^p}{|\boldsymbol{x} - \boldsymbol{y}|^{n+sp}} d\boldsymbol{x} \, d\boldsymbol{y} \right)^{\frac{1}{p}} . \tag{3.2}$$

We can also define the so-called Gagliardo semi-norm as

$$[u]_{W^{s,p}(\Omega)} = \left(\int_{\Omega} \int_{\Omega} \frac{|u(\boldsymbol{x}) - u(\boldsymbol{y})|^p}{|\boldsymbol{x} - \boldsymbol{y}|^{n+sp}} d\boldsymbol{x} \, d\boldsymbol{y} \right)^{\frac{1}{p}} \tag{3.3}$$

This definition of the fractional Sobolev space cannot be extended to the case $s \geq 1$. In literature there are some relationship to describe the behavior of the Gagliardo semi-norm for $s \to 0$ or $s \to 1$ [13]. For $u \in W^{1,p}(\Omega)$ holds

$$\lim_{x \to 1^-} (1 - s) \int_{\Omega} \int_{\Omega} \frac{|u(\boldsymbol{x}) - u(\boldsymbol{y})|^p}{|\boldsymbol{x} - \boldsymbol{y}|^{n+sp}} d\boldsymbol{x} \, d\boldsymbol{y} = C' \int_{\Omega} |\nabla u|^p d\boldsymbol{x} , \tag{3.4}$$

for a suitable positive constant C' depending only on n and p. Similarly, by extending the functional space of u to $W^{s,p}(\mathbb{R}^n)$, it can be demonstrated that

$$\lim_{x \to 0^+} s \int_{\Omega} \int_{\Omega} \frac{|u(\boldsymbol{x}) - u(\boldsymbol{y})|^p}{|\boldsymbol{x} - \boldsymbol{y}|^{n+sp}} d\boldsymbol{x} \, d\boldsymbol{y} = C'' \int_{\Omega} |u|^p d\boldsymbol{x} , \tag{3.5}$$

for a suitable positive constant C'' depending only on n and p [88]. For $p = 2$, it can be seen that the limit (3.4) goes to the $H^1(\Omega)$ norm multiplied by the constant C', and the limit (3.5) goes to the $L^2(\mathbb{R}^n)$ norm multiplied by the constant C''.

Proposition 3.1. *Let $0 < s \leq s_1 < 1$ and $u : \Omega \to \mathbb{R}$ be a measurable function. Then*

$$\|u\|_{W^{s,p}(\Omega)} \leq C \|u\|_{W^{s_1,p}(\Omega)} , \tag{3.6}$$

for some suitable positive constant $C = C(n, s, p) \geq 1$. In particular,

$$W^{1,p}(\Omega) \subset W^{s_1,p}(\Omega) \subseteq W^{s,p}(\Omega) . \tag{3.7}$$

The first relationship ($W^{1,p}(\Omega) \subset W^{s_1,p}(\Omega)$) it's effective under regularity hypotheses on $\partial\Omega$.

In this work we will only consider the case $p = 2$, so that $W^{s,p}(\Omega) = H^s(\Omega)$ is an Hilbert space. In such case, as said before, it holds

$$W^{1,p}(\Omega) = H^1(\Omega) \subset H^s(\Omega) . \tag{3.8}$$

The relationship (3.8) will be considered later to show that the search for the regularization term in the boundary optimal control theory occurs in a narrow subset (the $H^1(\Gamma)$ space) of the correct functional space ($H^{1/2}(\Gamma)$). From (3.1) we define the functional space $H^s(\Omega)$ as

$$H^s(\Omega) = \left\{ u \in L^2(\Omega) : \frac{|u(\boldsymbol{x}) - u(\boldsymbol{y})|}{|\boldsymbol{x} - \boldsymbol{y}|^{\frac{n}{2}+s}} \in L^2(\Omega \times \Omega) \right\} . \qquad (3.9)$$

3.1.2 The fractional Laplacian as a nonlocal operator

As mentioned before, in this work the only case with $p = 2$ is considered. This introduce the $H^s(\mathbb{R}^n) = W^{s,2}(\mathbb{R}^n)$ space. This class of functional spaces turns out to be important since they are Hilbert spaces. They are related to the fractional Laplacian $(-\Delta)^s$, defined as

$$(-\Delta)^s u(\boldsymbol{x}) = C(n,s) \ P.V. \int_{\mathbb{R}^n} \frac{u(\boldsymbol{x}) - u(\boldsymbol{y})}{|\boldsymbol{x} - \boldsymbol{y}|^{n+2s}} d\boldsymbol{y}, \qquad 0 < s < 1 , \qquad (3.10)$$

where $P.V.$ means "in the principle value sense" [13]. $C(n,s)$ is a constant that depends on s and on the dimensionality of the problem n. Its value can be written as

$$C(n,s) = s \ 2^{2s} \ \frac{\Gamma(\frac{n+2s}{2})}{\pi^{n/2} \ \Gamma(1-s)} , \qquad (3.11)$$

where $\Gamma(\cdot)$ is the Gamma function, defined as $\Gamma(k) = \int_0^\infty t^{k-1}e^{-t}dt$.

The integral over $\boldsymbol{y}$ in equation (3.10) suggests that the local behavior (on a generic point $\boldsymbol{x}$) of the operator $(-\Delta)^s$ is influenced by the values of the function over all the considered domain. This causes that $(-\Delta)^s$ can be considered a *nonlocal operator*. The fractional diffusion models can be thought of as being a special case of the general nonlocal diffusion models. In general, the numerical simulation of nonlocal operators is non-trivial, and the same goes for the fractional Laplacian simulations.

Nonlocal operators

Nonlocal models have been recently used in many applications, including continuum mechanics [89], kinetic equations, phase transitions [90], nonlocal heat conduction, turbulence models [91], etc. In particular, such models become very important when "local" classical models are not able to represent the physical phenomenon.

In the classical partial differential equation models interaction between two domains occur only due to contact. In the nonlocal models interaction can occur at a distance. In particular, let $\Omega \in \mathbb{R}^n$ be a bounded, open domain. So we can define the nonlocal diffusion operator $\mathfrak{L}$ applied on the function $u(\boldsymbol{x}) : \Omega \to \mathbb{R}$ as

$$\mathfrak{L}u(\boldsymbol{x}) = 2 \int_{\mathbb{R}^n} (u(\boldsymbol{y}) - u(\boldsymbol{x}))\gamma(\boldsymbol{x}, \boldsymbol{y})d\boldsymbol{y} , \qquad \forall \boldsymbol{x} \in \Omega \subseteq \mathbb{R}^n , \qquad (3.12)$$

where the kernel $\gamma(\boldsymbol{x}, \boldsymbol{y}) : \Omega \times \Omega \to \mathbb{R}$ is a non-negative, symmetric mapping [15]. Now we consider the steady-state nonlocal diffusion

$$\begin{cases} -\mathfrak{L}u(\boldsymbol{x}) = f(\boldsymbol{x}) & \text{on } \Omega , \\ u(\boldsymbol{x}) = 0 & \text{on } \Omega_l . \end{cases} \qquad (3.13)$$

In this set of problems the boundary conditions are imposed on an interaction volume Ω_l, which is disjoint from Ω. In numerical analysis usually Ω_l is a finite domain that surround Ω. In particular, a domain Ω_l on a ball of radius δ is typically used. In such cases, the domain $\Omega_l = \Omega_{I_\delta}$ is defined as

$$\Omega_{I_\delta} = \{ \; \boldsymbol{y} \in \mathbb{R}^n \smallsetminus \Omega \text{ such that } \boldsymbol{y} \in B_\delta(x) \text{ for some } x \in \Omega \; \} , \qquad (3.14)$$

where $B_\delta(x)$ denotes the ball of radius δ centered at x.

Fractional Laplacian as a special case of the nonlocal Laplacian

Previously, fractional Laplacian and nonlocal operators have been introduced. By comparing the equation (3.10) with the equation (3.12) it can be seen that the fractional Laplacian is a special case of the nonlocal operators. In fact, if the kernel $\gamma(\boldsymbol{x}, \boldsymbol{y})$ in (3.12) is defined as

$$\gamma(\boldsymbol{x}, \boldsymbol{y}) = \frac{c_{n,s}}{2|\boldsymbol{y} - \boldsymbol{x}|^{n+2s}} \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n . \qquad (3.15)$$

Then the nonlocal operator can be written as

$$- \mathfrak{L}u(\boldsymbol{x}) = (-\Delta)^s, \quad 0 < s < 1 . \qquad (3.16)$$

In the following, we consider the system (3.13), where $\mathfrak{L}$ is defined as in (3.12), the kernel is defined as in (3.15) and $f \in L^2(\Omega)$. Under these hypotheses, the weak formulation of the fractional diffusive problem can be written as

$$\frac{c_{n,s}}{2} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{u(\boldsymbol{y}) - u(\boldsymbol{x})}{|\boldsymbol{x} - \boldsymbol{y}|^{n+2s}} \big(v(\boldsymbol{y}) - v(\boldsymbol{x})\big) \; d\boldsymbol{y} \; d\boldsymbol{x} = \int_\Omega fv \; d\boldsymbol{x} , \quad \forall v \in H_\Omega^s(\mathbb{R}^n) . \qquad (3.17)$$

The nonlocal models are more general than their fractional and local counterparts. In fact, the nonlocal models can be a connection between fractional and local models. In particular, the partial differential equations (PDEs) can be seen as the local limit ($\delta \to 0$) and fractional PDEs as the global limit ($\delta \to \infty$) of nonlocal models [92]. The restriction of the nonlocal model on the domain Ω is obtained imposing

$$u(\boldsymbol{x}) = 0, \ v(\boldsymbol{x}) = 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^n \smallsetminus \Omega \ . \tag{3.18}$$

3.1.3   Numerical modeling of Riesz fractional Laplacian

Several methods have been developed to perform the numerical simulations of the fractional Laplacian on bounded domains, and many works have been recently published on this topic [17, 93, 94]. In this work, the numerical simulation of fractional Laplacian with homogeneous Dirichlet boundary conditions is considered.

We introduce a numerical method for the modeling of the equation (3.17). Such a model is called Riesz fractional Laplacian [14]. In the following section, a numerical approach to the fractional Laplacian based on a spectral method is presented [17]. All the presented models are validated through a comparison with some benchmark results found in the literature.

The Riesz fractional Laplacian on a bounded domain Ω is based on the application of the real space formula (3.17) to functions defined in the considered domain. In order to restrict such equation to bounded domains, it is important to consider the condition (3.18). The integrals over $\mathbb{R}^n$ can be split as

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} (.) \ d\boldsymbol{y} d\boldsymbol{x} = \int_{\Omega} \int_{\Omega} (.) \ d\boldsymbol{y} d\boldsymbol{x} + \int_{\mathbb{R}^n \smallsetminus \Omega} \int_{\Omega} (.) \ d\boldsymbol{y} d\boldsymbol{x} +$$
$$+ \int_{\Omega} \int_{\mathbb{R}^n \smallsetminus \Omega} (.) \ d\boldsymbol{y} d\boldsymbol{x} + \int_{\mathbb{R}^n \smallsetminus \Omega} \int_{\mathbb{R}^n \smallsetminus \Omega} (.) \ d\boldsymbol{y} d\boldsymbol{x} \ . \tag{3.19}$$

Considering (3.18), it is straightforward to demonstrate that

$$\frac{c_{n,s}}{2} \int_{\mathbb{R}^n \smallsetminus \Omega} \int_{\mathbb{R}^n \smallsetminus \Omega} \frac{u(\boldsymbol{y}) - u(\boldsymbol{x})}{|\boldsymbol{x} - \boldsymbol{y}|^{n+2s}} \big(v(\boldsymbol{y}) - v(\boldsymbol{x})\big) d\boldsymbol{y} d\boldsymbol{x} = 0 \ . \tag{3.20}$$

We impose now the kernel $\gamma(\boldsymbol{x}, \boldsymbol{y}) = \frac{v(\boldsymbol{y}) - v(\boldsymbol{x})}{|\boldsymbol{x} - \boldsymbol{y}|^{n+2s}}$. It is also possible to derive

the following

$$\int_{\mathbb{R}^n \smallsetminus \Omega} \int_\Omega \gamma(\boldsymbol{x}, \boldsymbol{y})\big(v(\boldsymbol{y}) - v(\boldsymbol{x})\big) d\boldsymbol{y} d\boldsymbol{x} =$$
$$= \int_\Omega \int_{\mathbb{R}^n \smallsetminus \Omega} \gamma(\boldsymbol{x}, \boldsymbol{y})\big(v(\boldsymbol{y}) - v(\boldsymbol{x})\big) d\boldsymbol{y} d\boldsymbol{x} \tag{3.21}$$
$$= \int_\Omega u(\boldsymbol{x}) v(\boldsymbol{x}) \int_{\mathbb{R}^n \smallsetminus \Omega} \frac{1}{|\boldsymbol{x} - \boldsymbol{y}|^{n+2s}} \, d\boldsymbol{y} \, d\boldsymbol{x} \,.$$

Then the fractional Laplacian for bounded domains turns to

$$\frac{c_{n,s}}{2} \int_\Omega \int_\Omega \frac{u(\boldsymbol{y}) - u(\boldsymbol{x})}{|\boldsymbol{x} - \boldsymbol{y}|^{n+2s}}\big(v(\boldsymbol{y}) - v(\boldsymbol{x})\big) \, d\boldsymbol{y} d\boldsymbol{x} +$$
$$+ c_{n,s} \int_\Omega u(\boldsymbol{x}) v(\boldsymbol{x}) \int_{\mathbb{R}^n \smallsetminus \Omega} \frac{1}{|\boldsymbol{x} - \boldsymbol{y}|^{n+2s}} \, d\boldsymbol{y} d\boldsymbol{x} = \int_\Omega f v \, d\boldsymbol{x} \,, \tag{3.22}$$

for all $v \in H_\Omega^s(\mathbb{R}^n)$.

In literature, many works are based on the extension of the bounded domain by an auxiliary domain Ω_l in order to model the mixed integral over $\mathbb{R}^n \smallsetminus \Omega$ in (3.22). In such works, usually the nonlocal behavior of the fractional Laplacian is calculated inside a ball $B_\lambda(\boldsymbol{x})$ of radius λ around the point $\boldsymbol{x}$ [15]. In these cases Ω_l is defined as $\bigcup_i B_\lambda(\boldsymbol{x}_i)$, $\forall \boldsymbol{x}_i \in \Omega$. So the integral over Ω turns out to be defined over $\Omega \cup \Omega_l$, and the integral over $\mathbb{R}^n \smallsetminus \Omega$ turns out to be defined over $(\mathbb{R}^n \smallsetminus \Omega) \cap B_\lambda(\boldsymbol{x}_i)$. In this work, a different approach will be used, with a semi-analytical method applied to the mixed integrals.

Finite element approximation for the Riesz method - I

In this section, the finite element approximation of Riesz bounded fractional Laplacian with homogeneous Dirichlet boundary conditions is introduced. In particular, we will focus on the first double integral in equation (3.22). Such integral is discretized as

$$\frac{c_{n,s}}{2} \int_\Omega \int_\Omega \frac{u(\boldsymbol{y}) - u(\boldsymbol{x})}{|\boldsymbol{x} - \boldsymbol{y}|^{n+2s}}\big(v(\boldsymbol{y}) - v(\boldsymbol{x})\big) \, d\boldsymbol{y} d\boldsymbol{x} \simeq$$
$$\simeq \frac{c_{n,s}}{2} \sum_{el_x=1}^{N_{el}} \sum_{x_g=0}^{ng(el_x)} J_{x_g} w_{x_g} \left(\sum_{el_y=1}^{N_{el}} \sum_{y_g=0}^{ng(el_y)} J_{y_g} w_{y_g} \frac{u(\boldsymbol{y}_g) - u(\boldsymbol{x}_g)}{|\boldsymbol{x}_g - \boldsymbol{y}_g|^{n+2s}}\big(v(\boldsymbol{y}_g) - v(\boldsymbol{x}_g)\big) \right). \tag{3.23}$$

The implementation of this model for a finite element solver is quite challenging. First of all, the kernel $\gamma(\boldsymbol{x}_g, \boldsymbol{y}_g)$ can be singular when $el_x = el_y$.

In order to avoid the singularity, different quadrature rules for the $\boldsymbol{x}$ and $\boldsymbol{y}$ discretizations can be used.

However, in this work, a trivial adaptive technique has been used in order to avoid the singularity, keeping at the same time a good approximation of the kernel near the singularities. This technique consists of the subdivision of the y-cell in 2^n sub-cells that share $\boldsymbol{x}_g$ as a node. The nodes that don't match the el_x nodes are calculated with an interpolation technique, so this method is developed for polygonal cells. Moreover the algorithm has been developed to perform multiple levels of refinement. The four sub-cells are divided in further 2^n sub-cells each, using the midpoint of each edge of the considered cell. In figure 3.1 the refinement on a single cell is shown. Under
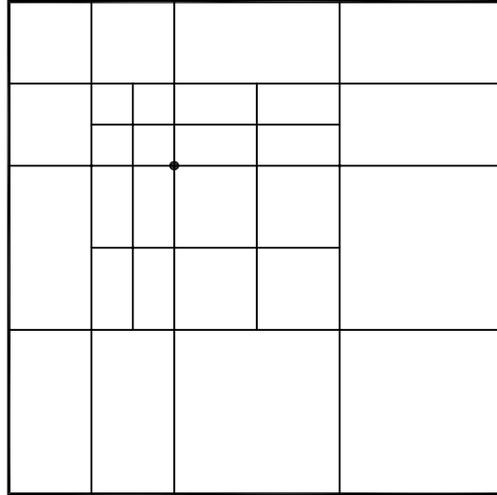
Figure 3.1: Refined cell for a generic quadrature point $\boldsymbol{x}_g$ (2 levels).

these hypotheses, the summation $\sum_{y_g=0}^{n_g(el_y)}$ turns out to be

$$\sum_{y_g=0}^{n_g(el_y)} \cdot = \sum_{n=0}^{N_{RefCells}(N_{split})} \sum_{y_g=0}^{n_g(el_y(n))} \cdot \quad , \tag{3.24}$$

where $N_{RefCells}$ is the number of refined cells obtained from N_{split} splits. The considered quadrature node $\boldsymbol{x}_g$ cannot be an internal point for any of the $N_{RefCells}$ (at most, can be a node for the cell). Since usually the quadrature points $\boldsymbol{y}_g$ of the cells cannot coincide with one of the nodes of the cell, the singularity of the kernel $\gamma(\boldsymbol{x}_g, \boldsymbol{y}_g)$ is avoided with this technique. It is also important to note that the refined cells accumulate towards the quadrature

point $\boldsymbol{x}_g$. This lead to a more precise numerical simulation of the behavior of the kernel near the singularity.

Since the Riesz approach is nonlocal, as reported above, the numerical system to be solved is characterized by a dense matrix. In particular, for implementation purposes, the matrix assembly should be divided into four different terms, so that the generic element of the stiffness matrix A_{ij} can be written as

$$A_{ij} = A_{ij}^{11} + A_{ij}^{12} + A_{ij}^{21} + A_{ij}^{22} \, ,$$

as shown in algorithm 4. However, as reported in [95], since the kernel is a symmetric function in $\boldsymbol{x}$ and $\boldsymbol{y}$, it holds $A_{ij}^{11} = A_{ij}^{22}$ and $A_{ij}^{12} = A_{ij}^{21}$. Thus, in the reported algorithm, we consider

$$A_{ij} = 2A_{ij}^{11} + 2A_{ij}^{12} \, .$$

Obviously, the resolution of a dense matrix is numerically expensive, so the resolution of these operators over complex domains might be unfeasible. PETSc [30] supports the resolution of numerical systems characterized by dense matrices, and in this work such an environment has been used.

In addition, it is important to underline that the cycle on the el_y in algorithm 4 it is defined, for $el_y = el_x$, over all the refined cells. The use of several levels of refinement allows building a coarser mesh, with a meaningful reduction of the computational cost of the simulation of fractional Laplacian. The most significant contribution to the fractional modeling consists precisely in the assembly for $el_y = el_x$ (near the singularity) and in the mixed term (the second integral in equation (3.22)). For this reason, the use of the refinement is strongly recommended in order to have better simulations with limited computational effort.

Finite element approximation for the Riesz method - II

In this section, the finite element approximation of the second integral of equation (3.22) will be introduced. Many works in literature present various techniques for the numerical simulation of the mixed integral. The most commonly used technique consists of the limitation of the nonlocal behavior of the fractional Laplacian to a ball B_δ of dimension n and radius δ built around the considered Gauss node $\boldsymbol{x}_g$ [12]. With such an approach the numerical domain must be extended outside of Ω, and the representation of the integral on $(\mathbb{R}^n \smallsetminus \Omega) \cap B_\delta$ is a numerical approximation of the mixed integral.

Algorithm 4 Local matrix assembly for Riesz fractional Laplacian.

---

1) Cycle on the processors, for $k_{proc} < n_{procs}$. The current processor is i_{proc}.

2) Double cycle on the elements el_y and el_x (equation (3.23)), for $el_{offset}[k_{proc}] \leq el_y < el_{offset}[k_{proc} + 1]$ and $el_{offset}[i_{proc}] \leq el_x < el_{offset}[i_{proc} + 1]$. Inside these cycles the mappings map_{el_y} and map_{el_x} are calculated.

3) Double cycle on the Gauss points x_g and y_g of the element el_x and el_y, respectively. Evaluation of the test functions $v(x_g) = v_1$ and $v(y_g) = v_2$ and of the jacobean and weights $w(\boldsymbol{x}_g) = w_{xg}$ and $w(\boldsymbol{y}_g) = w_{yg}$.

for $ig = 0, n_g(el_x)$ **do**

　　for $jg = 0, n_g(el_y)$ **do**

　　　　4) Double cycle on the test functions.

　　　　for $i = 0, nDof1$ **do**

　　　　　　for $j = 0, nDof2$ **do**

　　　　　　　　5) Assembly of the local matrix, splitted in four components

$$A_{ij}^{11} \mathrel{+}= \frac{C_{ns}\, v_1[j]\, v_1[i]\, w_{xg}\, w_{yg}\, J_{xg}\, J_{yg}}{\|\boldsymbol{x}_g - \boldsymbol{y}_g\|^{n+2s}}$$

$$A_{ij}^{12} \mathrel{+}= \frac{C_{ns}\, (-v_2[j])\, v_1[i]\, w_{xg}\, w_{yg}\, J_{xg}\, J_{yg}}{\|\boldsymbol{x}_g - \boldsymbol{y}_g\|^{n+2s}}$$

　　　　　　end for

　　　　end for

　　end for

end for

6) The assembled local matrices are mapped into the global matrix through the following mappings: $A^{11} \to map_{el_x}$ for rows and columns. $A^{12} \to map_{el_x}$ for rows and map_{el_y} for columns.

However, in this work, a different approach will be used. In [14], an approach based on the restriction of the mixed integral over $\mathbb{R}^n \smallsetminus \Omega$ to the boundary $\partial\Omega$ is introduced. A similar approach will be used in this work, based on both a numerical and an analytical approach to the mixed integral. Let consider the integral over $\mathbb{R}^n \smallsetminus \Omega$

$$\int_{\mathbb{R}^n \smallsetminus \Omega} \frac{1}{|\boldsymbol{x}_g - \boldsymbol{y}|^{n+2s}}\, d\boldsymbol{y}\,, \quad \boldsymbol{x} \in \Omega\,.$$

For a fixed $\boldsymbol{x}_g$, a loop on the entire boundary of the numerical problem ($\partial\Omega$) has been implemented. In particular, as shown in Figure 3.2, each boundary element is divided into a number of subdivisions determined by the user. Obviously, a greater number of subdivisions leads to a better nu-
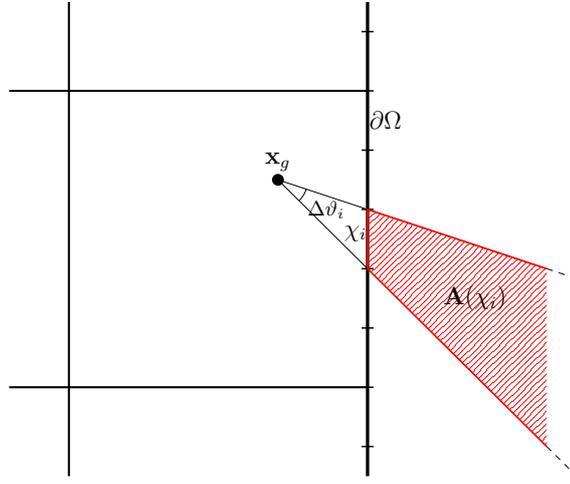
Figure 3.2: Area of numerical integration for one of the subdivisions of $\partial\Omega$.
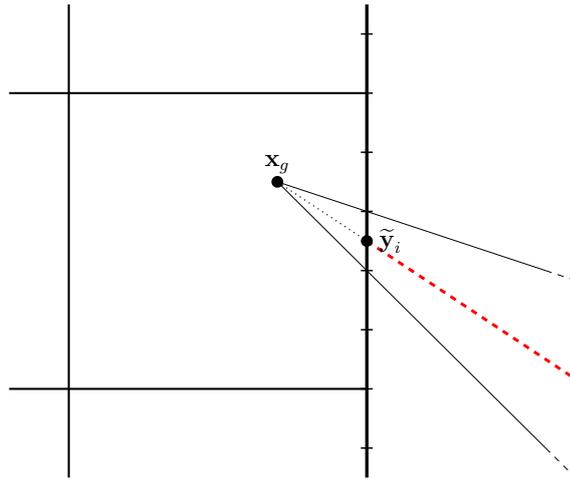
Figure 3.3: Numerical integration of the mixed integral for one of the subdivisions of $\partial\Omega$.

merical evaluation of the mixed integral, but to greater computational cost. Each subdivision χ_i represents an approximation of the area $A(\chi_i)$, such that $\bigcup_i A(\chi_i) = \{\mathbb{R}^n \setminus \Omega\}$. So we obtain

$$\fint_{\mathbb{R}^n \setminus \Omega} \frac{1}{|\boldsymbol{x}_g - \boldsymbol{y}|^{n+2s}} \, d\boldsymbol{y} = \sum_{i=1}^{N_{sub}} \int_{A(\chi_i)} \frac{1}{|\boldsymbol{x}_g - \boldsymbol{y}|^{n+2s}} \, d\boldsymbol{y} \ , \tag{3.25}$$

where N_{sub} is the total number of subdivision of the whole boundary $\partial\Omega$. The integrals over $A(\chi_i)$ are approximated along the line connecting $\boldsymbol{x}_g$ and $\widetilde{\boldsymbol{y}}_i$.

As can be seen in Figure 3.3, the node $\widetilde{\mathbf{y}}_i$ is the medium point of the segment χ_i. The integration 3.25 will be approximated along the line $\mathcal{I}$ connecting $\boldsymbol{x}_g$ and $\widetilde{\mathbf{y}}_i$. So considering t as the distance between $\boldsymbol{x}_g$ and $\boldsymbol{y} \in \mathcal{I}$, the approximated integral results as

$$\int_{A(\chi_i)} \frac{1}{|\boldsymbol{x}_g - \boldsymbol{y}|^{n+2s}} \; d\boldsymbol{y} \; \simeq \Delta\vartheta_i \int_{|\boldsymbol{x}_g - \widetilde{\mathbf{y}}_i|}^{+\infty} \frac{1}{t^{n+2s}} dt \; , \tag{3.26}$$

which implies

$$\begin{aligned}
\Delta\vartheta_i \int_{|\boldsymbol{x}_g - \widetilde{\mathbf{y}}_i|}^{+\infty} \frac{1}{t^{n+2s}} dt &= \Delta\vartheta_i \Big[\frac{1}{-(n+2s-1)t^{n+2s-1}} \Big]_{|\boldsymbol{x}_g - \widetilde{\mathbf{y}}_i|}^{+\infty} \\
&= \frac{\Delta\vartheta_i}{(n+2s-1)|\boldsymbol{x}_g - \widetilde{\mathbf{y}}_i|^{n+2s-1}} \; .
\end{aligned} \tag{3.27}$$

Therefore, the mixed integral of equation (3.22) turns out to be

$$\begin{aligned}
c_{n,s} \int_\Omega u(\boldsymbol{x}) v(\boldsymbol{x}) \int_{\mathbb{R}^n \smallsetminus \Omega} & \frac{1}{|\boldsymbol{x} - \boldsymbol{y}|^{n+2s}} \; d\boldsymbol{y} d\boldsymbol{x} \simeq \\
&\simeq \sum_{el_x=1}^{N_{el}} \sum_{x_g=0}^{n_g(el_x)} J_{x_g} w_{x_g} u(\boldsymbol{x}_g) v(\boldsymbol{x}_g) \sum_{i=1}^{N_{sub}} \frac{\Delta\vartheta_i}{(n+2s-1)|\boldsymbol{x}_g - \widetilde{\mathbf{y}}_i|^{n+2s-1}} \; .
\end{aligned} \tag{3.28}$$

Now the variational formulation for mixed integrals has been presented. The extension of it to one-dimensional domains is straightforward. It is important to underline that the refinement method introduced in the last paragraph can be extended at the mixed integrals, in order to have a more refined discretization of the boundary $\partial\Omega$.

The complete finite element model for Riesz fractional Laplacian is obtained by the sum of the discretized term in equation (3.23) and (3.28). The numerical modeling is completed by the discretized source term, $\int_\Omega f v d\boldsymbol{x}$, and by homogeneous boundary conditions as

$$\begin{aligned}
\sum_{el_x=1}^{N_{el}} \sum_{x_g=0}^{n_g(el_x)} & J_{x_g} w_{x_g} \left(\sum_{el_y=1}^{N_{el}} \sum_{y_g=0}^{n_g(el_y)} J_{y_g} w_{y_g} \frac{u(\boldsymbol{y}_g) - u(\boldsymbol{x}_g)}{|\boldsymbol{x}_g - \boldsymbol{y}_g|^{n+2s}} \big(v(\boldsymbol{y}_g) - v(\boldsymbol{x}_g) \big) \right) \\
&+ \sum_{el_x=1}^{N_{el}} \sum_{x_g=0}^{n_g(el_x)} J_{x_g} w_{x_g} u(\boldsymbol{x}_g) v(\boldsymbol{x}_g) \sum_{i=1}^{N_{sub}} \frac{\Delta\vartheta_i}{(n+2s-1)|\boldsymbol{x}_g - \widetilde{\mathbf{y}}_i|^{n+2s-1}} = \\
&= \sum_{el_x=1}^{N_{el}} \sum_{x_g=0}^{n_g(el_x)} J_{x_g} w_{x_g} f(\boldsymbol{x}_g) v(\boldsymbol{x}_g) \; .
\end{aligned}$$

$$\tag{3.29}$$

3.2 The Dunford-Taylor fractional Laplacian

In this section, a technique for the simulation of the fractional Laplacian based on the spectral theory is presented. When the spatial domain is bounded, the fractional powers can be defined in terms of the Fourier series, which leads to a conceptual simplification of a complex problem such as that of fractional operators. The spectral properties of the Laplacian will be extended to fractional Laplacian. The eigenvalue problems applied to PDE will be introduced at first. Then, models based on the Dunford-Taylor method for both spectral and integral fractional Laplacian will be presented. This section will be based on the Dunford-Taylor representation of fractional Laplacian defined as

$$(-\Delta)^{-s} = \frac{1}{2\pi i} \int_{\mathcal{D}} z^{-s}(z+\Delta)^{-1}dz, \qquad 0 < s < 1 \,. \tag{3.30}$$

This topic has been widely developed by Bonito et al., and the interested reader can see [96, 17, 16, 97].

3.2.1 The eigenvalues of the Laplacian

The eigenvalue problems applied to PDE are introduced in this section, with particular emphasis on the Laplacian eigenvalues. First of all, some basic concepts are introduced. The eigenvalue problem for the Laplacian can be written as

$$\begin{cases} \Delta f = \lambda f \\ B.C. \end{cases} \tag{3.31}$$

In this framework we introduce now some basic definitions and theorems.

Definition 3.3. *We define the Rayleigh quotient as*

$$R = \frac{\|\nabla z\|^2_{\mathcal{L}^2(\Omega)}}{\|z\|_{\mathcal{L}^2(\Omega)}} = \frac{\int_\Omega |\nabla z|^2\,dx}{\int_\Omega |z|^2\,dx} \,. \tag{3.32}$$

Now, considering $\mathcal{Y}$ as the space of the test function for a Dirichlet problem, the following theorem can be introduced.

Theorem 3.1. *Let $z \in \mathcal{Y}$ a minimum point for the Rayleigh quotient, so that*

$$m = \frac{\|\nabla z\|^2_{\mathcal{L}^2(\Omega)}}{\|z\|_{\mathcal{L}^2(\Omega)}} = min_{w \in \mathcal{Y}} \frac{\|\nabla w\|^2_{\mathcal{L}^2(\Omega)}}{\|w\|_{\mathcal{L}^2(\Omega)}} \,, \tag{3.33}$$

therefore m is the first eigenvalue of Δ_D, with the corresponding eigenfunction z. Δ_D is the Laplacian with a set of Dirichlet boundary conditions.

The Theorem 3.1 shows that at least one of the eigenvalues of the Laplacian can be found. Another more general theorem is now introduced to find all the Laplacian eigenvalues.

Theorem 3.2. *Let $v_1, v_2, \cdots, v_{n-1}$ be the first $n-1$ orthogonal eigenvectors of Δ_D. Let also consider*

$$\mathcal{Y}_n = \{w \in \mathcal{Y} : (w, v_i) = 0, \forall i = 1, \cdots, n-1\}.$$

If exists v_n such as the Rayleigh quotient m_n is minimized on $\mathcal{Y}_n$,

$$m_n = min_{w \in \mathcal{Y}_n} = \frac{\|\nabla w\|_{\mathcal{L}^2(\Omega)}^2}{\|w\|_{\mathcal{L}^2(\Omega)}},$$

then $m_n = \lambda_n$, where λ_n is the n-th eigenvalue of Δ_D. The corresponding associated eigenfunction is v_n.

The two enunciated theorems have an extremely relevant scope: thanks to them we can not only conclude that it is possible to determine the eigenvalues and eigenfunctions of the Laplacian operator, but we also have an explicit way to do it. Moreover, it is possible to demonstrate that the eigenfunctions of the Dirichlet problem are a complete orthonormal system on $\mathcal{L}^2(\Omega)$. Such a system can be used to develop solutions of the boundary problems considered in the Fourier series. The properties listed above lead to the following theorem.

Theorem 3.3. *Let $\{f_n\}_{n=1}^{+\infty}$ a sequence of orthonormal eigenvectors in the Hilbert space $\mathcal{H}$, and are $c_1, c_2, \cdots, c_N$ numerical coefficients. Then*

$$\left\| f - \sum_{n=1}^{N} c_n f_n \right\| \geq \left\| f - \sum_{n=1}^{N} (f, f_n) f_n \right\|. \tag{3.34}$$

We call $\lambda_n = (f, f_n)$ as Fourier coefficients of f, with the respect of $\{f_n\}_{n=1}^{+\infty}$.

The theorem states that in approximating f with a partial sum, the minimum that occurs is the case where the approximation is given in the Fourier series development. Moreover, the Theorem 3.3 lead to the conclusion that $\forall f \in \mathcal{H}$ the Fourier coefficient series converges, however it is not possible to demonstrate a priori that converges to $\|f\|$. So we introduce the following proposition.

Proposition 3.2. *If the sequence $\{f_n\}_{n=1}^{+\infty}$ is a complete orthonormal system, the following Parseval equality applies*

$$\|f\|^2 = \sum_{n=1}^{+\infty} |(f, f_n)|^2 \,, \tag{3.35}$$

where the internal product (f, f_n) defines the n-th Fourier coefficient of f with an orthonormal base.

Lastly, using the Fourier transform we can obtain the following definition of fractional Sobolev space, which will be used in the following

$$H^s(\Omega) = \left\{ w \in L^2(\Omega) : \|w\|_{H^s(\Omega)} = \left(\int_\Omega (1 + |\xi|^2)^{\frac{s}{2}} |\mathscr{F}(w)(\xi)|^2 d\xi \right)^{\frac{1}{2}} < \infty \right\}, \tag{3.36}$$

with $w \in H^s(\Omega)$.

The theory on the eigenvalue problems and on the Fourier series presented above is now used to develop a numerical approach to the fractional Laplacian.

3.2.2 The spectral method

For functions defined over $\mathbb{R}^d$, there is a natural way to define the fractional Laplacian as a pseudo-differential operator using the Fourier transform $\mathscr{F}$. Given a function u in the Schwartz class $\mathscr{S}$, we have

$$(-\Delta)^s u = \mathscr{F}^{-1}\left(|\xi|^{2s} \mathscr{F} u\right). \tag{3.37}$$

The theory on the Laplacian eigenvalues introduced above is now lead back to the fractional Laplacian $(-\Delta)^s$. The Laplacian operator can be written as $-\Delta : \mathcal{D}(-\Delta) \to L^2(\Omega)$, where the domain $\mathcal{D}(-\Delta) = H_0^1(\Omega) \subset H^2(\Omega)$ is dense. The operator $-\Delta$ is unbounded, positive and closed, and its inverse is compact. Therefore, there exists a countable collection of eigenpairs $\{\lambda_k, \phi_k\}_{k \in \mathbb{N}} \subset \mathbb{R}^+ \times H_0^1(\Omega)$ such that $\{\phi_k\}_{k \in \mathbb{N}}$ is an orthonormal basis of $L^2(\Omega)$ as well as an orthogonal basis of $H^0(\Omega)$ [17]. Thus, considering $u \in \{v \in L^2(\Omega) : (-\Delta)^s v \in L^2(\Omega)\}$, $0 < s < 1$, the fractional Dirichlet Laplacian can be defined as

$$(-\Delta)^s u = \sum_{k=1}^{\infty} \lambda_k^s (u, \phi_k) \phi_k \,, \tag{3.38}$$

with $k \in \mathbb{N}$. As shown previously, $(.,.)$ represents the inner product, and in this framework it is defined on $L^2(\Omega)$. $\{\phi_k\}$ is an orthonormal basis of eigenfunctions of $-\Delta$ corresponding to eigenvalues $\{\lambda_j\}$. Now it is possible to redefine the fractional Sobolev space H^s in a spectral way as

$$H^s(\Omega) = \left\{ u = \sum_{k=1}^{\infty} (u, \phi_k)\phi_k \quad : \quad \sum_{k=1}^{\infty} \lambda_k^s (u, \phi_k)^2 < \infty \right\}. \qquad (3.39)$$

The spectral fractional Laplacian is based on a direct approximation of the inverse of the fractional operator. Then, instead of solving $(-\Delta)^s u = f$ we approximate $u = (-\Delta)^{-s} f$. It is important to underline that the spectral approach is a method to find the solution of $(-\Delta)^s u = f$, but it is not a numerical representation of the fractional Laplacian operator. In the next section, the integral method will be presented, in order to have the representation of the operator. The approximation of $u = (-\Delta)^{-s} f$ can be obtained through (3.30)

$$(-\Delta)^{-s} f = \frac{1}{2\pi i} \int_{\mathcal{D}} z^{-s} (z + \Delta)^{-1} f dz. \qquad (3.40)$$

Since the operator $-\Delta$ is positive, we can consider $\mathcal{D}$ so as to obtain the Balakrishnan formula [98]

$$(-\Delta)^{-s} f = \frac{\sin(s\pi)}{\pi} \int_0^{\infty} \mu^{-s} (\mu \boldsymbol{I} - \Delta)^{-1} f d\mu. \qquad (3.41)$$

The spectral method here presented is based on this formula. However, this approach isn't well defined for integral fractional Laplacian.

Now a change of variables in the integral of equation (3.41) is needed to develop a numerical method for the simulation of the spectral Laplacian. We put $\mu = e^{-2t}$, which implies $d\mu = -2e^{-2t}dt$. Since $t = -\frac{log\mu}{2}$, the new limits of integration are $t(\mu = 0) = +\infty$ and $t(\mu = +\infty) = -\infty$. Under these hypotheses we obtain

$$u = (-\Delta)^{-s} f = \frac{2\sin(s\pi)}{\pi} \int_{-\infty}^{+\infty} e^{2st} (\boldsymbol{I} - e^{2t}\Delta)^{-1} f dt. \qquad (3.42)$$

In the following $c_s = \frac{2\sin(s\pi)}{\pi}$ will be used. Now, a sinc quadrature will be used for the numerical representation of (3.42). The approximated value of u results

$$u \simeq u^k = c_s k \sum_{j=-N^-}^{N^+} e^{2st_j} (\boldsymbol{I} - e^{2t_j}\Delta)^{-1} f = c_s k \sum_{j=-N^-}^{N^+} e^{2st_j} w_j, \qquad (3.43)$$

where

$$w_j = (\boldsymbol{I} - e^{2t_j}\Delta)^{-1}f\,. \tag{3.44}$$

Moreover, k is the quadrature step size, $t_j = jk$ are the quadrature points, $N^+ = \lceil\frac{\pi^2}{4(1-s)k^2}\rceil$ and $N^- = \lceil\frac{\pi^2}{4sk^2}\rceil$, where $\lceil.\rceil$ denote the round up to the nearest integer. For further information on the sinc quadrature applied to the fractional model, one can see [16]. In particular, in the cited paper a complete error analysis of the used approximation is developed.

Now we can solve the equation (3.44) as a partial differential equation. We can apply the inverse operator $(\boldsymbol{I} - e^{2t_j}\Delta)^{-1}$ directly to w_j, obtaining the following PDE (in weak form)

$$(w_{h,j}, v_h)_\Omega + e^{2t_j}(\nabla w_{h,j}, \nabla v_h)_\Omega = (f, v_h)_\Omega\,, \quad \forall v_h \in V_h\,, \tag{3.45}$$

where $V_h \subset H_0^1(\Omega)$ is the finite element space consisting of continuous piecewise linear functions. The unknowns $w_{h,j}$ are defined over V_h for each j on the same mesh.

Implementation of spectral fractional Laplacian

The implementation of the spectral fractional Laplacian consists of the numerical simulation of $N^- + N^+ + 1$ finite element problems. Such problems are mutually independent. Therefore the parallel implementation is straightforward since every numerical problem can be solved by a different processor. The $N^- + N^+ + 1$ equations (3.45) only requires the implementation of a classical finite element solver for the diffusion-reaction problem. Then, once the solutions are found, an algebraic relation between the $w_{j,h}$ is needed in order to find the solution u, as provided from equation (3.43). Since the numerical problems are mutually independent, the solution matrix $\mathbf{A}$ can be split into $N^- + N^+ + 1$ matrices $\mathbf{A}_j$. The solution u^k can be found through a weighted sum of the $w_{j,h}$ on the Gauss nodes.

It is important to underline that the choice of k affects the approximation error and the computational cost. The lower the value of k, the better is the representation of the solution u^k. However, low k values lead to high N^- and N^+ values, therefore it is necessary to solve a large number of finite element problems. This could increase meaningfully the computational cost of the simulation.

The discrepancy between the exact solution and its approximation due to all the introduced approximations consist of the sinc quadrature error,

which is exponentially convergent, and the finite element approximation error, which is optimal up to a logarithmic factor depending on the smoothness of the data f. For further information on the error estimation, which is not part of this work, one can see [17, 16]. It is worth mentioning that the error analysis does not require the domain to be convex.

Lastly, it is important to note that for sufficiently high values of N^+ and N^- numerical catastrophic cancellation can happen, due to the nature of the exponential function. To avoid cancellations, one can consider $w' = e^{2st_j}w$ such that the equations (3.45) and (3.43) become, respectively

$$
\begin{aligned}
(w'_{h,j}, v_h)_\Omega + e^{2t_j}(\nabla w'_{h,j}, \nabla v_h)_\Omega &= e^{2st_j}(f, v_h)_\Omega \\
u^k = c_s k \sum_{j=-N^-}^{N^+} w'_j\,. &
\end{aligned}
\tag{3.46}
$$

The system of equation (3.46) has been implemented in a finite element based code, and tested on various 1D and 2D domains. With the presented change of variables, the numerical cancellations are avoided.

3.2.3   The integral method

The system (3.46) is a direct approximation of the inverse of the operator. This means that the spectral fractional Laplacian is not suitable for partial differential equations involving fractional Laplacians. The spectral method is useful to solve equations of the type $\Delta^s u = f$. In order to build properly a model for fractional Laplacian operator alternative to the Riesz exposed before, we apply the Parseval's equality (see Proposition 3.2) to equation $\Delta^s u = f$, finding the following weak formulation: given $f \in \mathbb{H}^{-s}(\Omega)$ find $u \in \mathbb{H}^s(\Omega)$ such that

$$
\int_\Omega |\xi|^s \mathscr{F}(u)|\xi|^s \overline{\mathscr{F}(w)}d\xi = (f, w)\,, \quad \forall w \in H^s\,,
\tag{3.47}
$$

where $\overline{\chi}$ is the complex conjugate of χ. Now, starting from equation (3.36) we can derive the following theorem.

Theorem 3.4. *Let* $s \in (0,1)$ *and* $0 \leq r \leq s$. *For* $u \in \mathbb{H}^{r+s}(\Omega)$ *and* $\theta \in \mathbb{H}^{s-r}(\Omega)$,

$$
\big((-\Delta)^{(s+r)/2}u, (-\Delta)^{(s-r)/2}\theta\big) = c_s \int_0^\infty t^{2-2s}\big(-\Delta(\boldsymbol{I} - t^2\Delta)^{-1}u, \theta\big)\frac{dt}{t}\,,
\tag{3.48}
$$

where

$$c_s = \left(\int_0^\infty \frac{y^{1-2s}}{1+y^2} dy \right)^{-1} = \frac{2\sin(s\pi)}{\pi} .$$

Proof. The proof of this theorem is taken from [97] and limited to the Ω domain. For more details, see the cited paper and references therein. Using the Parseval's theorem we have

$$(-\Delta(\boldsymbol{I} - t^2\Delta)^{-1}u, \theta) = \int_\Omega \frac{|\xi|^2}{1+t^2|\xi|^2} \mathscr{F}(u)(\xi)\overline{\mathscr{F}(\theta)(\xi)}d\xi . \tag{3.49}$$

Consequently, the right-hand side of equation (3.48) becomes

$$c_s \int_0^\infty t^{2-2s}(-\Delta(\boldsymbol{I} - t^2\Delta)^{-1}u, \theta)\frac{dt}{t} =$$

$$= c_s \int_0^\infty t^{1-2s} \int_\Omega \frac{|\xi|^2}{1+t^2|\xi|^2} \mathscr{F}(u)(\xi)\overline{\mathscr{F}(\theta)(\xi)}d\xi dt \tag{3.50}$$

In order to invoke Fubini's theorem, we now show that the right-hand side of equation (3.50) is equal to

$$c_s \int_\Omega |\mathscr{F}(u)(\xi)| \, |\mathscr{F}(\theta)(\xi)| \int_0^\infty t^{1-2s} \frac{|\xi|^2}{1+t^2|\xi|^2} dt d\xi =$$

$$= \int_\Omega |\xi|^{2s}|\mathscr{F}(u)(\xi)| \, |\mathscr{F}(\theta)(\xi)|d\xi , \tag{3.51}$$

which is finite for $u \in \mathbb{H}^r(\Omega)$ and $\theta \in \mathbb{H}^{s-r}(\Omega)$. We now apply Fubini's theorem and the change of variable $y = t|\xi|$ in (3.50) to obtain

$$\int_\Omega |\xi|^{2s}|\mathscr{F}(u)(\xi)| \, |\mathscr{F}(\theta)(\xi)|d\xi = \left((-\Delta)^{(s+r)/2}u, (-\Delta)^{(s-r)/2}\theta\right) . \tag{3.52}$$

This completes the proof of the theorem. $\qquad\square$

We can now consider the equation (3.48) as a first operator representation of the integral fractional Laplacian. Some new concepts are now introduced in order to have a numerical representation of it.

Similarly to the spectral fractional Laplacian, we define $w(\psi, \mu) = w(\mu) \in \mathbb{H}^1(\Omega)$ for a generic $\psi \in L^2(\Omega)$. The introduced $w(\mu)$ has to be the solution of

$$\int_\Omega w(\mu)\phi \, dx + \mu^2 \int_\Omega \nabla w(\mu) \cdot \nabla \phi \, dx = - \int_\Omega \psi\phi \, dx , \quad \forall \phi \in \mathbb{H}^1(\Omega) . \tag{3.53}$$

It is possible to demonstrate that the equation (3.48) is equivalent to

$$\int_\Omega |\xi|^{r+s}\mathscr{F}(\eta)(\xi)|\xi|^{s-r}\overline{\mathscr{F}(\eta)(\xi)}d\xi =$$
$$= \frac{2\sin(s\pi)}{\pi}\int_0^\infty \mu^{1-2s}\int_\Omega \big((-\Delta)(\boldsymbol{I}-\mu^2\Delta)^{-1}\eta\big)\theta\,dx\,d\mu\,. \tag{3.54}$$

Therefore, considering $u,\theta \in H^s(\Omega)$ with $s \in (0,1)$, the prior equation turns to

$$\int_\Omega |\xi|^{r+s}\mathscr{F}(\eta)(\xi)|\xi|^{s-r}\overline{\mathscr{F}(\eta)(\xi)}d\xi =$$
$$= \frac{2\sin(s\pi)}{\pi}\int_0^\infty \mu^{-1-2s}\left(\int_\Omega \big(u+w(u,\mu)\big)\theta\,dx\right)d\mu\,. \tag{3.55}$$

We can define now the integral fractional operator as the right-hand side of equation (3.55)

$$a(\eta,\theta) = \frac{2\sin(s\pi)}{\pi}\int_0^\infty \mu^{-1-2s}\left(\int_\Omega \big(\eta+w(\eta,\mu)\big)\theta\,dx\right)d\mu\,, \tag{3.56}$$

with $\eta,\theta \in H^s(\Omega)$. u has been substitute by η in order to show the generality of this operator. The operator $a(\eta,\theta)$ can be used for the numerical simulation of fractional Laplacian in PDEs. For now, we are interested in solving the simple equation (3.13), so we obtain

$$a(u,v) = (f,v)\,, \qquad \forall v \in H^s(\Omega)\,. \tag{3.57}$$

The equation (3.57) lead to ill-posed numerical systems. Usually, the system (3.57) is numerically solved with an extension of the variables $w(u,\mu)$ over the whole $\mathbb{R}^n$, with a subsequent reduction on a ball of a variable radius around every Gauss node. In this work, we will precondition the ill-posed system in order to have solutions on Ω, without involving larger domains.

The numerical simulation of the generic integral operator $a(\eta,v)$ is based, similarly to the spectral fractional Laplacian, on a sinc quadrature. We also perform a change of variable $\mu = e^{-\frac{1}{2}y}$ and we obtain

$$a(\eta,\theta) = \frac{\sin(s\pi)}{\pi}\int_{-\infty}^{+\infty} e^{sy}\left(\int_\Omega \big(\eta+w(\eta,\mu(y))\big)\theta\,dx\right)dy\,. \tag{3.58}$$

Given a quadrature spacing k, N^- defined as in the spectral case, $N^- = \lceil\frac{\pi^2}{4sk^2}\rceil$, and N^+ twice the corresponding spectral value, $N^+ = \lceil\frac{\pi^2}{2(1-s)k^2}\rceil$, the sinc quadrature is defined as

$$a^k(\eta,\theta) = \frac{\sin(s\pi)}{\pi}k\sum_{l=-N^-}^{N^+} e^{sy_l}\int_\Omega \big(\eta+w(\eta,\mu(y_l))\big)\theta\,dx\,. \tag{3.59}$$

We can model the numerical problem of $(-\Delta)^s u = f$ with the operator $a^k(\eta, \theta)$ and the equations (3.53) to determine the values of $w(\eta, \mu(y_l))$. The numerical model to be solved is

$$
\begin{aligned}
(a) \quad & a^k(u, \theta) = \frac{\sin(s\pi)}{\pi} k \sum_{l=-N^-}^{N^+} e^{sy_l} \big(u + w(u, e^{-\frac{1}{2}y_l}), \theta\big) = (f, \theta), \\
(b) \quad & (w_l, \phi) + e^{-y_l}(\nabla w_l, \nabla \phi) = -(u, \phi),
\end{aligned}
\tag{3.60}
$$

where $w_l = w(u, e^{-\frac{1}{2}y_l})$, $y_l = kl$, and $\forall \phi \in \mathbb{H}^1(\Omega)$, $\theta \in H^s(\Omega)$. The system of equations (3.60) is composed by $N^+ + N^- + 2$ equations. It should be noted that all the introduced functional spaces are defined over Ω, and so for the $L^2(\Omega)$ scalar products $(.,.)$. The implementation of the presented system is easy, however leads to ill-posed numerical systems. In the next paragraph, some measures for the implementation of the presented numerical problem will be introduced.

Implementation of integral fractional Laplacian

In contrast to the spectral fractional Laplacian, considering $N^* = N^+ + N^- + 1$, the $N^* + 1$ equations of system (3.60) are not mutually independent. Therefore it is impossible to split the resolution of the numerical problem into $N^* + 1$ separated numerical systems. The matrix corresponding to the system (3.60) will result as

$$
M = \left.\begin{bmatrix}
A_1 & 0 & \cdots & 0 & B \\
0 & A_2 & \cdots & 0 & B \\
\ddots & & \ddots & & \ddots \\
0 & 0 & \cdots & A_{N^*} & B \\
C_1 & C_2 & \cdots & C_{N^*} & D
\end{bmatrix}\right\}
\begin{matrix}
\\ \left.\vphantom{\begin{matrix}A\\0\\\ddots\\0\end{matrix}}\right\}(b) \\ \\ \left.\vphantom{C}\right\}(a)
\end{matrix}
$$

This behavior can be noticed printing the PETSc matrix corresponding to a generic one-dimensional simulation of the integral fractional Laplacian. The quadrature step size k has been chosen in order to have $N^+ = N^- = 10$ for $s = 0.5$. The resulting matrix m is reported in Figure 3.4. The light blue values in figure are zeroes, the blue and red are positive and negative non-zero values, respectively.

As pointed out above, usually this system is not solved in this form because of his ill-posedness. Many cited works (as [97] and references therein)
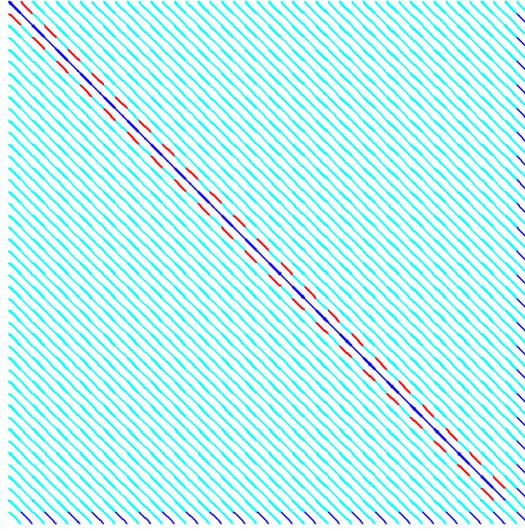
Figure 3.4: PETSc matrix representation for the one-dimensional case with $s = 0.5$ and $N^+ = N^- = 10$.

are based on the technique of the restriction on balls of a certain radius mentioned earlier. In this case, we used a preconditioners-based technique in order to solve the presented system defined over Ω.

First of all, a simple approach based on a combination of an additive and a multiplicative set of preconditioners has been used. In particular, the additive preconditioners have been used for the first N^* equations. The diagonal term of such equations corresponds to the unknown w_l. The additive preconditioners used for the solution of the w-equations imply that the first N^* equations are solved independently of each other using the old value of u. The multiplicative preconditioner used for the solution of the u-equation (the last one) implies that the u unknown is solved using the updated values of the w_l, solved by using the presented additive scheme.

We consider now a generic system of equation

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{B}^T & \mathcal{D} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} . \tag{3.61}$$

It is possible to trace back the numerical system of Figure 3.4 to the presented scheme, with $u_1 = w_l$ and $u_2 = u$ The numerical scheme of the "additive+multiplicative" method consists in

$$\begin{cases} u_1 = \mathcal{A}^{-1}(f - u_2^{old}\mathcal{B}) \\ u_2 = \mathcal{D}^{-1}(g - u_1\mathcal{B}^T) \, , \end{cases} \tag{3.62}$$

where in the second equation u_1 is obtained from the first equation. This combined additive and multiplicative technique converges to a solution. However, a study of the solver behavior showed that the convergence was reached through a very high number of sub iterations. This behavior could lead to high computational cost for the simulation of the fractional Laplacian in complex domains.

The numerical convergence of the integral fractional Laplacian can be improved with the Schur complement. We now consider the first line of (3.61), and we solve it as

$$u_1 = \mathcal{A}^{-1}(f_1 - \mathcal{B}^T u_2),$$

where the inverse of $\mathcal{A}$ is an approximation. In the following, we will refer to such approximated inverse as $(\mathcal{A}^*)^{-1}$. Now we substitute u_1 into the second line of (3.61)

$$-\mathcal{B}\mathcal{A}^{*-1}\mathcal{B}^T u_2 + \mathcal{D}u_2 = g - \mathcal{B}\mathcal{A}^{*-1}f,$$

obtaining

$$\underbrace{\left(\mathcal{D} - \mathcal{B}\mathcal{A}^{*-1}\mathcal{B}^T\right)}_{\text{Schur complement } \mathcal{S}} u_2 = g - \mathcal{B}\mathcal{A}^{*-1}f. \tag{3.63}$$

In order to find the solution u_2 it is necessary to invert the Schur complement as

$$\begin{cases} u_2 = \left(\mathcal{D} - \mathcal{B}\mathcal{A}^{*-1}\mathcal{B}^T\right)^{*-1} \left(g - u_1\mathcal{B}^T\right), \\ u_1 = \mathcal{A}^{-1}(f - u_2\mathcal{B}). \end{cases} \tag{3.64}$$

The inverse of $\mathcal{S}$ is approximated numerically through an incomplete LU factorization (ILU). Furthermore, some tests have shown that the number of iterations decreases strongly using a GMRES solver instead of the classic PREONLY implemented in the PETSc environment [31]. This becomes necessary when ill-conditioned systems are solved. In particular, in the case of system (3.60), and considering $u_1 = w_l$ and $u_2 = u$ as mentioned before, the ill-conditioned part of the numerical system lies in the $\mathcal{D}$ matrix inversion. The Schur complement avoid the $\mathcal{D}$ inversion, expecting the approximated inversion of $(\mathcal{D} - \mathcal{B}\mathcal{A}^{*-1}\mathcal{B}^T)$ that is not ill-conditioned. After the calculation of the inverse of $\mathcal{S}$, we can calculate u_2 and substitute it into the first equation for u_1.

The final used numerical scheme includes an additive scheme for the w_l calculations, and the Schur complement for the solution of u. The proposed method has been tested and needs only one iteration to reach the convergence.

We can conclude that the numerical performances are strongly improved with this technique.

3.3 Fractional Laplacian numerical results

In this section the results of the presented methods for the numerical simulation of the fractional Laplacian are presented. In recent years, many works have been published on this topic. In particular, we will refer to [14] as a benchmark for all the presented numerical techniques. The solutions are validated through comparison with other numerical simulations or through comparison with analytical solutions of the mathematical problem of the fractional Laplacian (only for one-dimensional simulations). All the presented simulations are performed on FEMuS, a Multiphysics Finite Element library with selective-adaptive mesh refinement and multigrid solvers developed at Department of Mathematics and Statistics of the Texas tech University [22].

3.3.1 One-dimensional simulations

In this section the solutions of the equation $(-\Delta)^s u = f$ will be presented for one-dimensional cases. We consider firstly the simple case $f = 1$ for the methods introduced in the last section: the Riesz, the spectral and the integral fractional Laplacian.

Riesz method

The system (3.13) can be solved analytically for the one-dimensional case. In particular we consider the analytic solution of such problem in the ball $B_R(0)$ of radius R in $\mathbb{R}^n$. It is possible to demonstrate [15] that for every function $f \in L^2(B_R(0))$ there exists a unique function $u \in H^s(\mathbb{R}^n)$ such as

$$u(x) = c_{n,s} \int_{B_R(0)} |x - y|^{2s-n} f(y) \int_0^{r_0(x,y)} \frac{r^{s-1}}{(r+1)^{n/2}} \, dr \, dy \,, \qquad (3.65)$$

with

$$c_{n,s} = s^{-2s} \frac{\Gamma(\frac{n}{2})}{R^2 \pi^{n/2} \Gamma(s)^2} \quad \text{and} \quad r_0(x,y) = \frac{(R^2 - |x|^2)(R^2 - |y|^2)}{|x - y|^2} \,. \qquad (3.66)$$

In the particular case of $f = 1$ in $B_1(x)$ we obtain

$$u(x) = 2^{-2s} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n+2s}{2})\Gamma(1+s)} (1 - x^2)^s \,. \qquad (3.67)$$

In Figure 3.5, the analytical solution obtained in equation (3.67) is compared with the solution of the solution of the Riesz model, presented in section 3.1. The simulations have been carried out on the one-dimensional domain $[-1, 1]$, with a coarse mesh of 2 elements and 6 refinement levels. Homogeneous Dirichlet boundary conditions have been applied and $N_{splits} = 4$ (see equation (3.24)) have been considered. As can be seen, the numerical results (u_R in the figure) match the analytical solution for all the tested values of s.
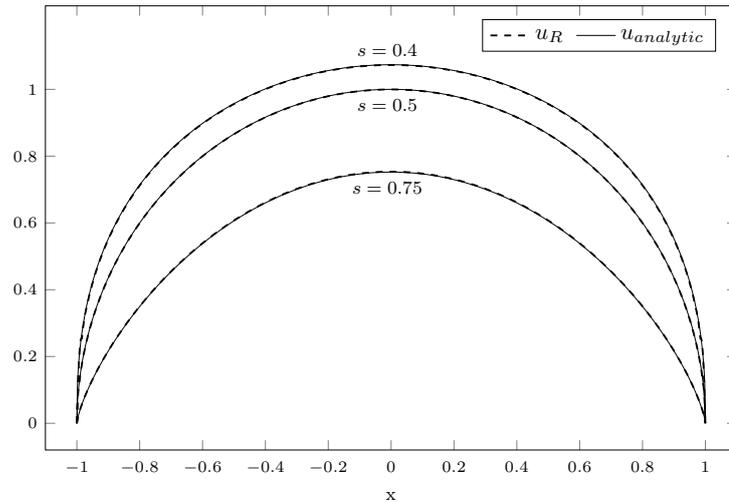
Figure 3.5: Comparison between the analytical and the Riesz solution of fractional Laplacian, for $s = 0.4, 0.5, 0.75$.

Table 3.1: L^2 norm of the solution u and comparison with the analytical solution for different refinement levels n_{lev} for $s = 0.5$. Note that $\| u_{analytical} \|_{L^2} = 1.154700$.

$n_{lev} (n_{dof})$	$\| u \|_{L^2}$	ε_{L^2}	$n_{lev} (n_{dof})$	$\| u \|_{L^2}$	ε_{L^2}
1 (5)	1.097736	0.049333	5 (65)	1.150361	0.003758
2 (9)	1.122600	0.027800	6 (129)	1.152515	0.001893
3 (17)	1.137947	0.014509	7 (257)	1.153603	0.000951
4 (33)	1.146131	0.007421	8 (513)	1.154151	0.000476

In table 3.1 the comparison between the L^2-norm of the analytical result and the numerical results of the $f = 1$ case are presented. It can be seen

that the solution u converges to the analytical solution as n_{lev} grows. The relative error ε_{L^2} have also been reported.

Spectral method

In Figure 3.6 the plot of the result of a numerical simulation with the spectral method on the same mesh used in the last example is reported, with $s = 0.5$ and a relatively high quadrature step size $k = 2$ that leads to $N^+ = N^- = 1$. To have better results, a minor k value should be considered. In this case, we have chosen that value in order to have a more compact representation in the figure.
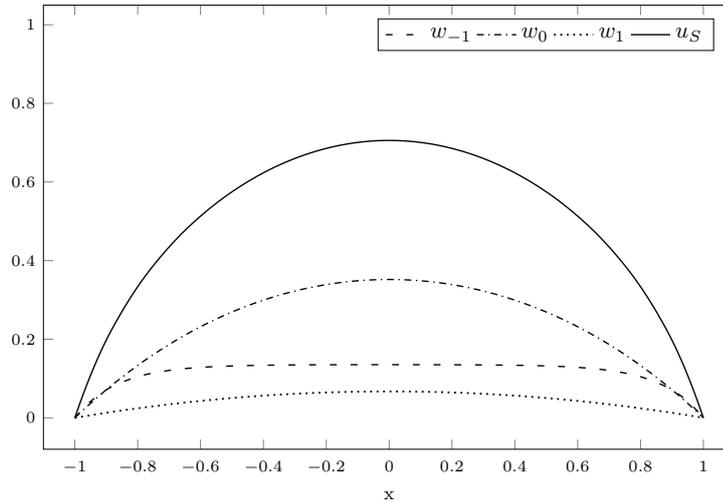
Figure 3.6: Plot of the unknowns u and all the w_j in a in a simple spectral fractional Laplacian case with $s = 0.5$, $k = 2$ and $f = 1$.

Note that the presented solution corresponds to the system (3.46), where a numerical catastrophic cancellation is avoided. The u value is proportional to the sum of the w_j. In Figure 3.6 one can see that the major contribution to the solution consists of the term w_0, and, for higher values of N^+ and N^-, it can be observed that w_j goes to zero when $j \to N^+$ or $j \to -N^-$.

From the comparison between Figures 3.5 and 3.6 it can be seen that the resulting u fields are not the same. This happens since the two implemented numerical models are not the same and represents two different kinds of fractional Laplacian, as can be seen in the benchmark in [14].

Integral method

The integral fractional Laplacian simulations are similar to the spectral ones. From a theoretical point of view, the two techniques solve the same approach to the fractional Laplacian (the Dunford Taylor based methods). However, while the integral method allows us to model the fractional operator, the spectral method is a direct approximation of the inverse.
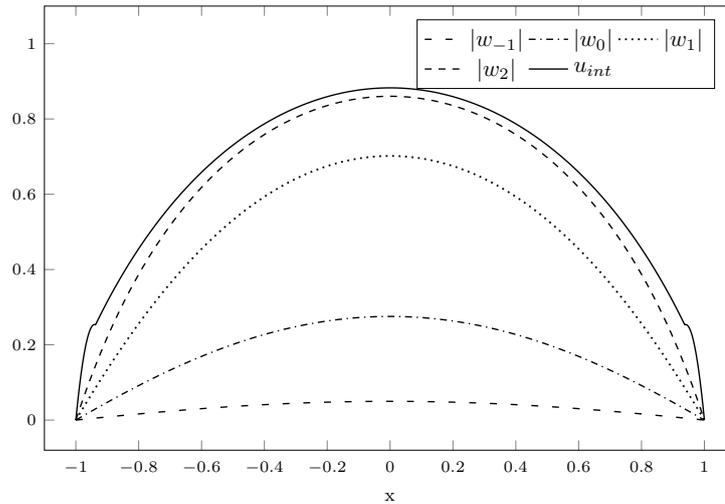
Figure 3.7: Plot of the unknowns u and all the absolute value of w_l, in a simple integral fractional Laplacian case with $s = 0.5$, $k = 2$ and $f = 1$.

In Figure 3.7 the result of the simulation of the same case $((-\Delta)^s u = 1)$ on the same mesh are reported. Again, relatively high quadrature step size $k = 2$ is considered, in order to have a limited number of unknowns w_l. The value of k leads to $N^+ = 2$ and $N^- = 1$. The values of the variables w_l are negative due to their definition (3.60), in Figure $|w_l|$ are reported. The u field reported in the Figure present some irregularities caused by the low value of k.

The u field obtained with the integral technique should be the same as the field obtained with the spectral technique. Again, the u results in Figure 3.7 and 3.6 doesn't match due to the low k value. Later will be shown that the two solutions converge for lower values of k. The presented simulation has been performed with the additive preconditioner system on w_l and with the Schurr complement on the u variable.

Comparison between methods

There is a structural difference between the Riesz fractional Laplacian and the methods based on the Dunford-Taylor technique. In a recent review work [14], whose purpose is to develop a numerical benchmark for the fractional Laplacian, it can be seen that the two numerical results are significantly different from each other. However, since the spectral and the integral fractional Laplacian are based on a similar mathematical model, the two results are expected to be similar.
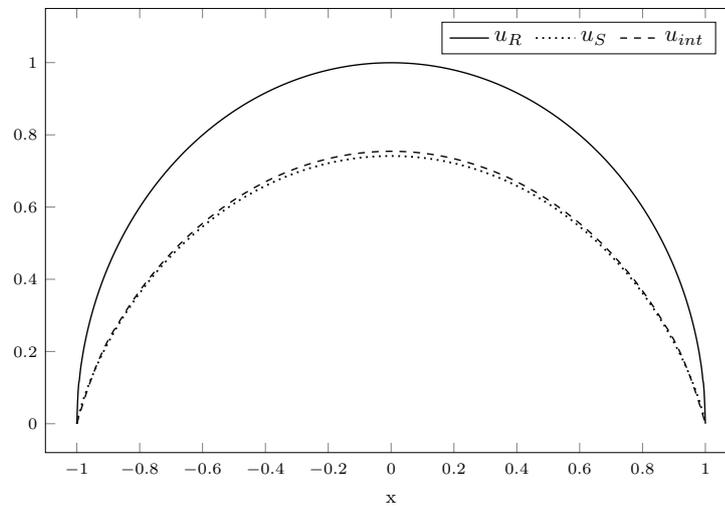
Figure 3.8: Comparison of the Riesz (u_R), spectral (u_S) and integral (u_{int}) solutions of $(-\Delta)^s = 1$ for $s = 0.5$.

In Figure 3.8 the numerical results for $s = 0.5$ are showed. In the following we will refer with u_R, u_S and u_{int} to the solutions of the Riesz, the spectral and the integral fractional Laplacian, respectively. In Figures 3.9, the results for $s = 0.4$ and $s = 0.75$ are showed. As can be seen higher values of s lead to lower values of u, because a stronger diffusive term is applied. In particular, for $s \to 1$, the solution of the models matches the solution of a Laplacian ($\Delta u = 1$).

It is particularly interesting to study the behavior of the solution of the models for $s \to 0$. In fact, in such case the equation of the fractional Laplacian can be traced back to the weak formulation of the mass matrix $(u, v) = (f, v)$, for all v in a consistent functional space. The solution of such an equation is $u = f$ and, in this particular case, $u = 1$. As can be seen in Figure 3.10, the three solutions tend to the correct solution. The numerical oscillations of
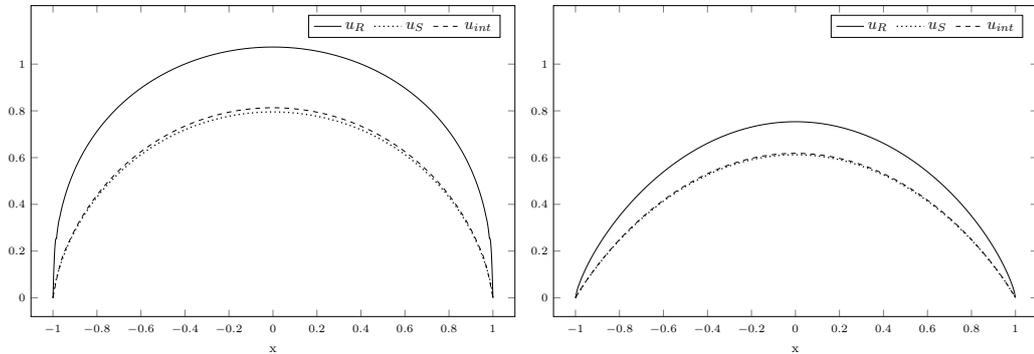
Figure 3.9: Comparison between the presented cases with $f = 1$ and $s = 0.4$ (left) and $s = 0.75$ (right).
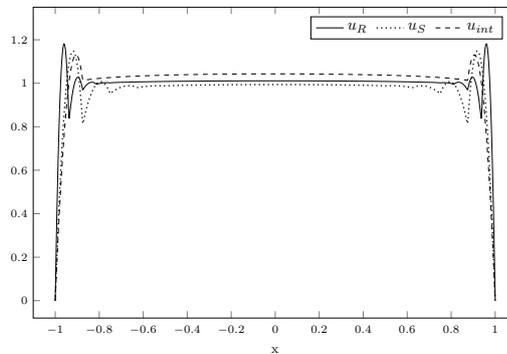
Figure 3.10: Comparison between the presented cases with $f = 1$ and $s = 0.01$. The solutions converge to the solution of a mass matrix.

the solutions near the boundaries are due to the imposition of homogeneous Dirichlet boundary conditions.

The sinusoidal source term

The presented methods have been tested with a sinusoidal source term. The simulated numerical problem is $(-\Delta)^s u = \sin(\pi x)$. These simulations aim to test all the presented algorithms with a source term dependent on x. In Figure 3.11 the results with the Riesz, spectral and integral methods are presented. The notation is the same used previously. The plot regards three different values of s, in particular $s = 0.25$, $s = 0.5$ and $s = 0.75$, and follow the previous cases with $f = 1$, with the spectral and the integral cases that present a similar behavior. The Riesz method shows slightly different values of the solution.
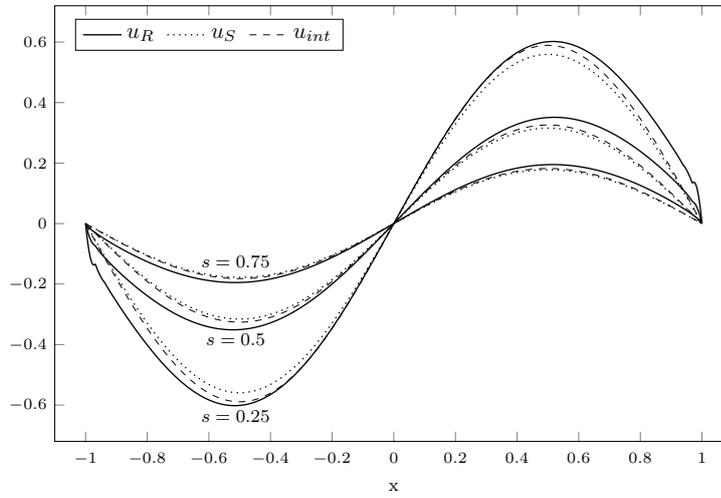
Figure 3.11: One-dimensional fractional Laplacian with sinusoidal source term for $s = 0.25$, $s = 0.5$ and $s = 0.75$, using the Riesz (u_R), spectral (u_S) and integral (u_{int}) methods.
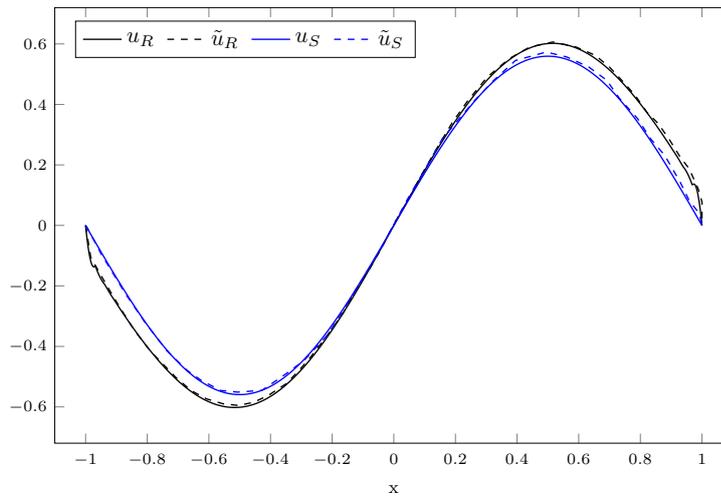
Figure 3.12: Comparison between the results of the numerical problem $(-\Delta)^{0.25}u = \sin(\pi x)$ between the implemented algorithm (solid lines) and the results taken from the cited review paper (dashed lines, $\tilde{u}_R$ and $\tilde{u}_S$). The Riesz case (black lines) and the spectral case (blue lines) are presented.

All the presented cases have been compared to the results of some review papers. In particular, in Figure 3.12 a comparison between the results obtained with our algorithm (solid line) and the numerical results presented in [14] (dashed lines, and represented as $\tilde{u}$). In particular, in black are presented

the results of the Riesz method and in blue the results of the spectral method. As can be seen in Figure, the results of the implemented algorithm match the numerical results presented in the literature. This is further evidence of the proper implementation of the presented algorithm for the one-dimensional simulations.

3.3.2 Two-dimensional simulations

In order to validate the developed algorithms, some tests on two-dimensional geometries have been carried out. In particular, a 2D simple square $\Omega = ([-1, 1], [-1, 1])$ have been considered. The mesh is a 2×2 mesh with a variable number of multigrid levels. In particular, 5 levels will be considered in the presented results. All the presented simulations include homogeneous Dirichlet boundary conditions. The numerical resolution of all the two-dimensional problems have been tested with different grid refinements, showing always good convergence properties.
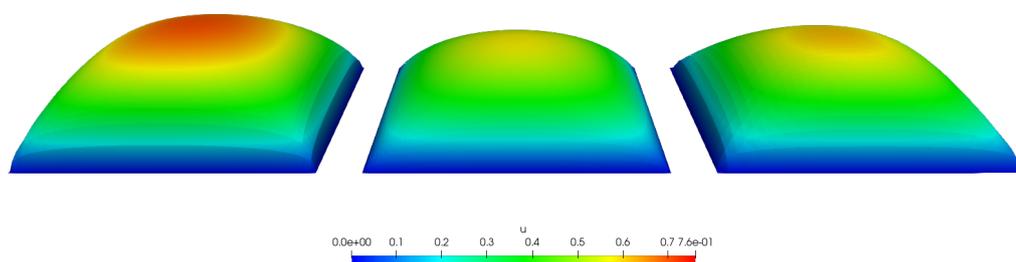
Figure 3.13: Comparison of the 2D results of the fractional Laplacian on a square with $f = 1$ and $s = 0.5$. From left to right: Riesz, spectral and integral method.

Constant source term

In Figure 3.13 the u solutions of the numerical problem $(-\Delta)^s u = 1$ on Ω are presented. The three curves present a similar behavior compared to the one-dimensional simulations. In Figure 3.14 the values of the unknown u along the line $x = 0$ are presented for $s = 0.5$. Again, it can be noted that the Riesz algorithm leads to higher values of the solution. The spectral and integral solutions are similar to each other. The two-dimensional fractional Laplacian
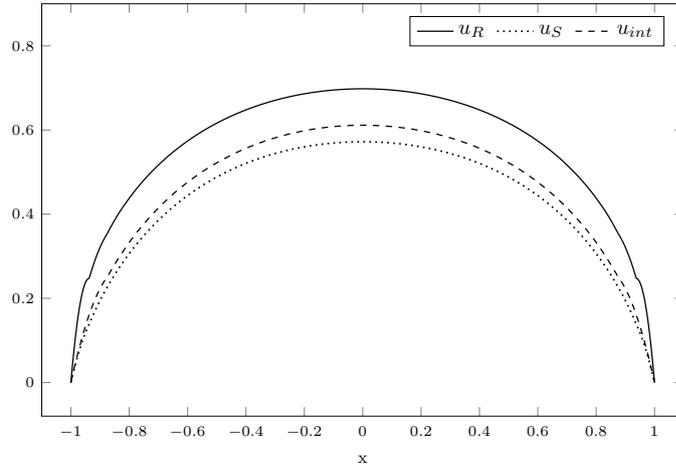
Figure 3.14: Comparison of the u field on the line $x = 0$ in the cases of Figure 3.13, with the Riesz (u_R), spectral (u_S) and integral (u_{int}) method.

will be used, in the framework of the optimal control problems, to model the regularization term on the boundaries of a three-dimensional simulation. For this reason, the three-dimensional results will not be reported in this work.
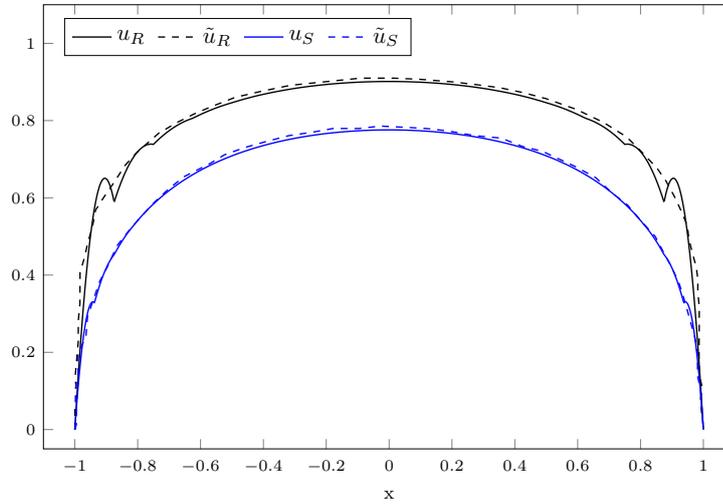
Figure 3.15:   Comparison between the results of the numerical two-dimensional problem $(-\Delta)^{0.25} u = 1$ between the implemented algorithm (solid lines) and the results taken from the literature (dashed lines, $\tilde{u}_R$ and $\tilde{u}_S$). The Riesz case (black lines) and the spectral case (blue lines) are reported.

In Figure 3.15, the comparison between the obtained results and the re-

sults taken from the literature is reported. In particular, the results used as a reference are taken from the same paper reported above. Likewise the previous comparison, our results are reported with solid lines, and the literature results with dashed lines. For both Riesz (black lines) and spectral (blue lines) methods, it can be noted that the behavior of the implemented algorithm matches the literature results.

Sinusoidal source term

Lastly, the sinusoidal source term case for two-dimensional simulation has been studied in order to test the algorithm with a variable source term. In Figure 3.16, the qualitative profiles of the solutions of the equation

$$(-\Delta)^s u = \sin(\pi x) \sin(\pi y) \tag{3.68}$$

has been reported. Note that the x and y used in this section differs from the same variables used in section 3.1. In fact, in that section the $\boldsymbol{x}$ and $\boldsymbol{y}$ variables were vector quantities representing the position in different integrals, in this case x and y represents the scalar position along one direction. In Figure it can be seen that the profiles of the solutions of the fractional Laplacian with a sinusoidal source term are consistent with all the results presented in literature.
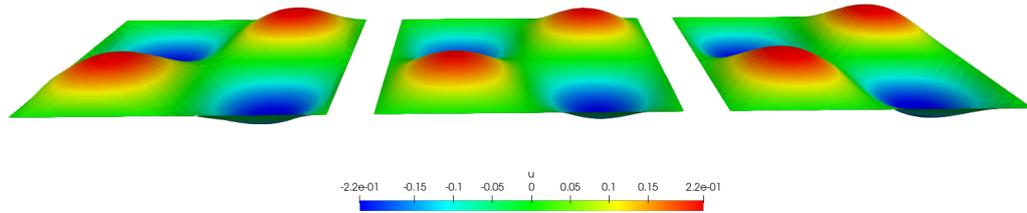
Figure 3.16: Comparison of the 2D results of the fractional Laplacian on a square with $f = \sin(\pi x) \sin(\pi y)$ and $s = 0.5$. From left to right: Riesz, spectral and integral method.

In Figure 3.17 the comparison between the results obtained with the implemented algorithm and the reference literature case is shown. Again, the numerical results of the implemented code match the literature results, both for the spectral and the Riesz algorithms. In this case, as in the literature results, the three methods show similar behavior. However, it is important
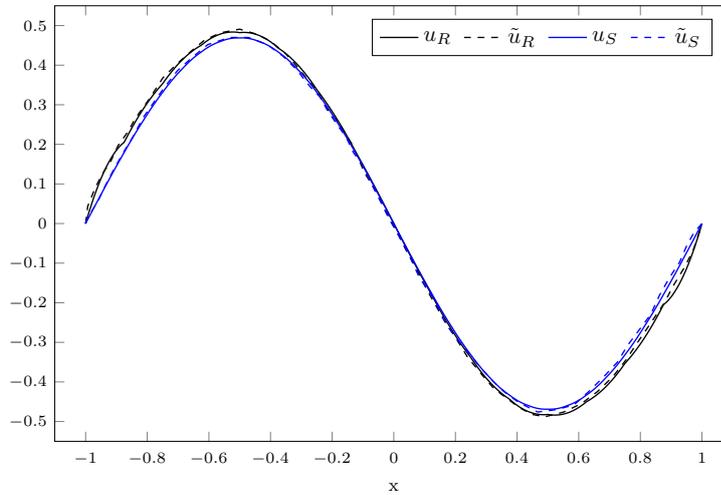
Figure 3.17:   Comparison between the results of the numerical two-dimensional problem (3.68) between the implemented algorithm (solid lines) and the results taken from the literature (dashed lines, $\tilde{u}_R$ and $\tilde{u}_S$). The Riesz case (black lines) and the spectral case (blue lines) are reported.

to remember that the simulation results of the fractional operators depend on the used method.

All the presented results are a validation of the implemented codes, as all the tested cases reproduce correctly the results taken from the literature. Since the model for the fractional Laplacian have been successfully validated, it is possible to use it to model the regularization in optimal control simulations.

3.4   Fractional operators in boundary optimal control problems

In this section, we consider again the optimal control problems introduced in chapter 2. As mentioned before, in the framework of such problems, the variables of a mathematical model are first divided into two classes: state and control variables. Then, an optimization problem is posed with respect to a certain objective functional of the above variables. Since the mathematical models we are interested in consist of Partial Differential Equations (PDE), we refer to this field as PDE-constrained optimal control.

Boundary optimal control problems are perhaps the most interesting class

of optimal control problems constrained by partial differential equations. In fact, the possibility of controlling the behavior of a physical system may often take place only by changing the values of certain quantities at the boundary of the domain. This is especially true if the interior of the physical system is not accessible or no physical mechanism can be triggered inside the domain from the outside. For instance, the temperature control of a physical system may be achieved either by optimizing boundary temperature values or boundary heat fluxes.

The nonlinear mathematical models that are constraints of boundary optimal control problems pose important theoretical and computational challenges and an increasing interest in research. Among them, we turn our attention to the occurrence of fractional derivatives. Theoretical results in mathematical analysis imply that the connection between functions defined on the domain of a PDE and their restriction to the boundary gives rise to fractional order Sobolev spaces. The presence of these spaces induces the necessity to deal with fractional derivatives and in particular with the very challenging problem of their numerical discretization, as we have seen in the last sections.

In this Section, we describe the occurrence of fractional norms in the treatment of boundary optimal control problems. These norms lead to the presence of fractional derivatives in the first-order necessary conditions that characterize optimal solutions. We recall that these first-order necessary conditions give rise to a system of equations often referred to as optimality system or *Karush-Kuhn-Tucker* (KKT) system. The formulation of boundary control problems constrained by PDEs can be characterized by the property of balanced regularity that will be described in this section.

Note that, although many works in literature concerning the optimal control with fractional operators have been published, most of them study the fractional Laplacian as constraint equation. Only a few works analyze the issue of balanced regularity by using fractional operators for the regularization term. This can be considered as a novelty introduced in this work.

3.4.1   The issue of balanced regularity

We introduce now the property of balanced regularity that naturally arises when dealing with boundary optimal control problems in function spaces. Only to keep the exposition simple, we describe these features with a model

problem given by a Laplace constraint operator, Dirichlet boundary conditions, and a tracking-type cost functional. In this work no further extension of the fractional operators to complex physics (i.e. the Koiter shell equation presented in the previous chapter) is developed. However, the aim of this work is the future development of this technique in order to apply the fractional Laplacian to boundary optimal control of complex systems. We remark that the issues we will be highlighting are also encountered in boundary optimal control problems with general PDE constraints, boundary conditions other than Dirichlet (such as Neumann or Robin), and cost functionals of non-tracking type.

Let us first recall some basic definitions for the symbols that will be used in the following. Let Ω be a bounded domain with boundary an and unit normal vector field n. For any domain $\mathcal{O} \subseteq \Omega$, we denote with $H^m(\mathcal{O})$ the classical Sobolev spaces. For any $\Gamma \subset \partial\mathcal{O}$ we denote with

$$\gamma_0 \,:\, H^m(\mathcal{O}) \to H^{m-1/2}(\Gamma) \tag{3.69}$$

the trace operator on Γ. This operator is surjective [43]. Also, for any boundary portion $\Gamma \subseteq \partial\mathcal{O}$, we use the notation $H^1(\mathcal{O}; \Gamma)$ for functions $H^1(\mathcal{O})$ with zero trace on Γ. When $\Gamma = \partial\mathcal{O}$, we also use the notation $H^1(\mathcal{O}) = H^1(\mathcal{O}; \partial\mathcal{O})$. The symbols ∇_Γ and Δ_Γ denote the surface gradient and surface Laplacian on $\partial\Omega$, respectively.

A model boundary optimal control problem

In order to describe the issue of balanced regularity, we introduce now a class of simple Dirichlet boundary optimal control problem. We denote the control boundary as $\Gamma_c \subseteq \partial\Omega$. On Γ_c, Dirichlet optimal conditions are sought. We recall now the problem 2.1 presented in the chapter 2.

Problem 3.1. *Find a state-control pair $(u, q) \in H^1(\Omega) \times Q$ which minimizes the cost functional*

$$\mathcal{J}(u, q) = \frac{1}{2}\|u - u_d\|^2_{L^2(\Omega_d)} + \frac{\alpha}{2}G(q)^2 \,, \tag{3.70}$$

under the constraints

$$\begin{align}
(\nabla u, \nabla v) &= 0 & &\forall v \in H^1_0(\Omega) \,, \tag{3.71}\\
u &= q & &\text{on } \Gamma_c \,, \tag{3.72}\\
u &= 0 & &\text{on } \partial\Omega \smallsetminus \Gamma_c \,. \tag{3.73}
\end{align}$$

Now we properly model the cost functional $G(q)$. The regularization constant $\alpha > 0$ is given and guarantees the existence of a minimizer of the presented optimal control problem. The desired state function u_d is also given. Moreover, Q is a Hilbert space on the domain Γ_c, depending on the Dirichlet boundary conditions in (3.72) and on the choice of the functional $G(q)$ in (3.70).

Note that the boundary condition (3.72) is intended in the sense of traces, meaning that it should more precisely be written using the trace operator as $\gamma_0 u = q$ on Γ_c. This immediately shows that, while $H^1(\Omega)$ is the space for the optimal states, the natural space where optimal boundary controls q should be sought is $Q = H^{1/2}(\Gamma_c)$, as dictated by the range of the trace operator (3.69). Thus, we can introduce the following definition.

Definition 3.4. *A boundary optimal control problem has the* balanced regularity property *if the optimal states belong to Sobolev spaces H^α with differentiability index α, while the boundary Dirichlet optimal controls belong to Sobolev spaces $H^\beta(\Gamma)$ with differentiability index $\beta = \alpha - \frac{1}{2}$, as dictated by the range of the trace operator.*

Therefore, it is the choice of the particular form of $G(q)$ which is responsible for guaranteeing the balanced regularity property of boundary control.

Usually, $G(q)$ is modeled as a $H^1(\Gamma_c)$ norm. However, this approach features the drawback given by a more restrictive control space than the natural $H^{1/2}$, that is dictated by the range of the trace operator. In literature, many works have been introduced to overcome the described lack of balanced regularity, without involving fractional norms. These approaches reformulate the problem based on the concept of lifting functions. The lifting function method is based on the addition of an auxiliary domain Ω_c such that $\Gamma_c \subseteq \partial\Omega_c$. Ω_c can be internal (when $\Omega_c \subseteq \Omega$) or external (when $\Omega_c \cap \Omega = \emptyset$). Under these hypotheses, $G(q)^2 = \|\tilde{q}\|_{H^1(\Omega_c)}^2$, where $\tilde{q}$ is the restriction of q to the domain Ω_c. The interested reader on the lifting function approach can see [49, 87, 99, 100].

The $H^{1/2}$ approach

The most natural choice of $G(q)$ that guarantees the existence of a boundary optimal control in $H^{1/2}(\Gamma_c)$ is given by

$$G(q) = \|q\|_{H^{1/2}(\Gamma_c)}. \tag{3.74}$$

In fact, when $u \in H^1(\Omega)$, the most natural choice for the regularization term on the boundary is exactly $H^{1/2}(\Gamma_c)$.

In literature many works used this approach to regularize a large variety of functionals in different frameworks. For example, a fractional regularization term has been applied to control problems with adaptive wavelet schemes [101] or with an energy space finite element approach [102].

Under the presented hypotheses, we can reformulate the optimal control problem 3.1.

Problem 3.2. *Find a state-control pair $(u, q) \in H^1(\Omega) \times Q$ which minimizes the cost functional*

$$\mathcal{J}(u, q) = \frac{1}{2}\|u - u_d\|^2_{L^2(\Omega_d)} + \frac{\alpha}{2}\|q\|^2_{H^{1/2}(\Gamma_c)}, \qquad (3.75)$$

under the constraints

$$(\nabla u, \nabla v) = 0 \qquad\qquad \forall v \in H^1_0(\Omega), \qquad (3.76)$$

$$u = q \qquad\qquad on\ \Gamma_c, \qquad (3.77)$$

$$u = 0 \qquad\qquad on\ \partial\Omega \smallsetminus \Gamma_c. \qquad (3.78)$$

Thanks to the choice of $G(q)$, the presented problem satisfies the balanced regularity property. The differentiation of the fractional norm on q induces an optimality system that contains the fractional derivative of q. Usually, alternative optimal control techniques are implemented in order to circumvent the presence of fractional derivatives. However, thanks to the numerical model introduced in the previous sections of this chapter, it is possible to simulate the fractional derivatives of q.

3.4.2   The optimality system

The main difference between the presented problem and problem 2.1 presented in the previous chapter is the regularization term of the functional $\mathcal{J}$. In order to obtain the optimality system in this case we consider the Gagliardo semi-norm obtaining

$$\mathcal{J}(u, q) = \frac{1}{2}\|u - u_d\|^2_{L^2(\Omega_d)} + \frac{\alpha}{2} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{\left(q(x) - q(y)\right)^2}{|x - y|^{n+1}} \, dy \, dx, \qquad (3.79)$$

where the reduction to Γ_c is obtained by imposing

$$q = \begin{cases} 0 & \text{on } \mathbb{R}^n \smallsetminus \Gamma_c, \\ q & \text{on } \Gamma_c, \end{cases} \qquad (3.80)$$

and n is the dimensionality of Γ_c (in this case $n = 1$). Therefore the derivative of $\mathcal{J}$ reads

$$\mathcal{J}' = \int_{\Omega_d} \delta u(u - u_d)\, d\boldsymbol{x} + \alpha \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\big(q(x) - q(y)\big)\big(\delta q(x) - \delta q(y)\big)}{|x - y|^2}\, dy\, dx\,. \quad (3.81)$$

It can be noted that, following a procedure similar to section 2.1.2, the optimality condition in this case reduces to

$$\alpha \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\big(q(x) - q(y)\big)\big(\delta q(x) - \delta q(y)\big)}{|x - y|^2}\, dy\, dx - \int_{\Gamma_c} \nabla \lambda \cdot \boldsymbol{n} \delta q\, dx = 0\,, \quad (3.82)$$

where λ is the Lagrange multiplier introduced. We can consider δq as a test function v. It is easy to show that the first term of (3.82) is the fractional Laplacian of q bounded to Γ_c under the hypotheses (3.80). We can now introduce the discretized optimality system based on the fractional Laplacian with $s = 0.5$.

The discretized optimality system

Some considerations about the block structure of the introduced optimality system are now introduced, together with some simple preconditioning techniques. It is important to note that, in contrast with the last chapter, the optimality system in this section is solved with a monolithic solver for u, q and λ. Under this condition, the optimality system arising from the first-order necessary conditions is a coupled system of differential equations whose numerical solution is quite challenging. Generally, the numerical system is nonlinear. Its non-linearity is originated by nonlinear PDE constraints and non-quadratic cost functionals.

Despite the complexity of the introduced numerical problem, the fully coupled solution of the optimality system is the most robust option for the numerical computation of candidate optimal states and controls. To this end, the attention focuses on preconditioned Krylov subspace methods who prove to be the most attractive solvers for linear algebraic systems arising from linearization of nonlinear problems.

The optimality system arising from the model Problem 3.2 consists of an algebraic linear system that exhibits a block structure. A finite element discretization of such an optimality system gives rise to a block structure of

the type

$$
\begin{bmatrix}
M_{\Omega_d} & 0 & -\Delta_\Omega + \left.\dfrac{\partial}{\partial n}\right|_{\Gamma_c} \\[2ex]
0 & \alpha F_{\Gamma_c} & -\left.\dfrac{\partial}{\partial n}\right|_{\Gamma_c} \\[2ex]
-\Delta_\Omega & 0 & 0
\end{bmatrix}
\begin{bmatrix}
u_h \\[2ex] q_h \\[2ex] \lambda_h
\end{bmatrix}
=
\begin{bmatrix}
u_d \\[2ex] 0 \\[2ex] 0
\end{bmatrix},
\tag{3.83}
$$

where u_h, q_h and λ_h denote the discretize state, control and adjoint variables, respectively. The blocks $M_\mathcal{O}$, $-\Delta_\mathcal{O}$ and $\left.\dfrac{\partial}{\partial n}\right|_\mathcal{O}$ denote the mass matrix, the Laplacian matrix and boundary Neumann matrix with integration over the generic domain $\mathcal{O}$, respectively. However, in this context the most interesting term in (3.83) is $F_\mathcal{O}$, which represents the matrix corresponding to the numerical implementation of the fractional Laplacian. As we have seen in the last section, usually $F_\mathcal{O}$ turns out to be a dense matrix.

In this work, the Riesz method is used for the numerical simulation of fractional Laplacian. Since the boundary of a two-dimensional case is a one-dimensional domain, the mixed integral for the Riesz assembly can be reduced to

$$
\begin{aligned}
c_{n,s} \int_\Omega \int_{\mathbb{R}^n \smallsetminus \Omega} & \frac{q(y) - q(x)}{|x - y|^{n+2s}} \big(v(y) - v(x) \big)\ dy dx = \\
&= c_{1,s} \int_\Omega q(x) v(x) \int_{\mathbb{R} \smallsetminus \Omega} \frac{1}{|x - y|^{1+2s}}\ dy dx \\
&= c_{1,s} \int_\Omega q(x) v(x) \left(\int_{-\infty}^{e_1} \frac{1}{(x - y)^{1+2s}}\ dy + \int_{e_2}^{\infty} \frac{1}{(x - y)^{1+2s}}\ dy \right) dx \\
&= \frac{c_{1,s}}{2s} \int_\Omega q(x) v(x) \left(\frac{1}{(x - e_2)^{2s}} - \frac{1}{(x - e_1)^{2s}} \right) dx ,
\end{aligned}
\tag{3.84}
$$

where e_1 and e_2 are the extremes of the segment Γ_c. Then, as we have seen in the Algorithm 4 and, in particular, in equation (3.29), the discretized form of the fractional Laplacian for a generic element $el_x \in \Gamma_c$ reads

$$
\begin{aligned}
\sum_{x_g=0}^{n_g(el_x)} & \frac{c_{1,s}}{2} J_{x_g} w_{x_g} \left(\sum_{el_y=1}^{N_{el}} \sum_{y_g=0}^{n_g(el_y)} J_{y_g} w_{y_g} \frac{q_h(y_g) - q_h(x_g)}{|x_g - y_g|^{n+2s}} \big(v(y_g) - v(x_g) \big) \right) \\
& + \sum_{x_g=0}^{n_g(el_x)} \frac{c_{1,s}}{2s} J_{x_g} w_{x_g} q_h(x_g) v(x_g) \left(\frac{1}{(x_g - e_2)^{2s}} - \frac{1}{(x_g - e_1)^{2s}} \right),
\end{aligned}
\tag{3.85}
$$

where all the loops are developed on the boundary elements. In fact, all the values in (3.85) refers to boundary values, e.g. $n_g(el_x)$ is the boundary

number of Gauss nodes, N_{el} the number of elements of Γ_c, etc. As we have seen in algorithm 4, in this case the assembly of the matrix should be done by splitting the assembly in four different contributions, F_{11}, F_{12}, F_{12} and F_{22}. Since $F_{11} = F_{22}$ and $F_{12} = F_{21}$ as reported in section 3.1, we can consider only the F_{11} and F_{12} contributions as $F_{\Gamma_c} = 2F_{11} + 2F_{12}$. Therefore, the local assembly on a Gauss node x_g (integral of el_x) and y_g (integral of el_{yc}) can be written as

$$F_{11}(i,j) = \frac{C_{1,s}}{2} w_{xg} w_{yg} \frac{v_1[j]\,v_1[i]}{\|x_g - y_g\|^{1+2s}} + \frac{C_{1,s}}{2s} w_{xg} \left(\frac{v_1[i]}{(x_g - e_2)^{2s}} - \frac{v_1[i]}{(x_g - e_1)^{2s}} \right),$$
$$F_{12}(i,j) = \frac{C_{1,s}}{2} w_{xg} w_{yg} \frac{(-v_2[j])\,v_1[i]}{\|x_g - y_g\|^{1+2s}} + \frac{C_{1,s}}{2s} w_{xg} \left(\frac{v_1[i]}{(x_g - e_2)^{2s}} - \frac{v_1[i]}{(x_g - e_1)^{2s}} \right).$$

Under these hypotheses, the sub-matrix F_{Γ_c} is clearly dense, with a sparsity pattern determined by the number of the degree of freedom of Γ_c.

The system (3.83) can be solved directly, in contrast to the method introduced in the last section where an iterative algorithm has been developed in order to find the optimal solution.

The inequality constraints

The study of optimal control problems is appealing for the transfer to technological applications. Thus, since these applications deals often with inequality constraints, it is important to consider the addition of such constraints to the PDE of the optimal control problems. Both state and control variables may be subject to inequality constraints. In some cases, inequality constraints may need to be enforced not because of practical considerations but due to theoretical conditions, e.g. to guarantee well-posedness. Thus, inequality constraints play an important role, and at the same time they make the analysis as well as the numerical discretization more challenging.

This is valid in particular for boundary optimal control problems. The problem 3.2 can be completed with inequality constraints. For instance, if we consider the case of box inequality constraints on the considered control variable problem, it can be restated together with inequality of the type

$$q_a \leq q \leq q_b \text{ a.e. on } \Gamma_c,$$

where $q_a, q_b \in L^\infty(\Gamma_c)$. The numerical realization of these constraints can be achieved using the *Primal-Dual Active Set* (PDAS) method. This method was first proposed by K. Ito, K. Kunisch, and M Bergounioux [103]. Each

iteration of the method consists in the solution of a linear system $\mathcal{M}_k \delta \boldsymbol{u}_{k+1} = f(\boldsymbol{u}_k)$, where $\mathcal{M}_k$ can be write as a 4×4 block structure

$$
\mathcal{M}_k = \begin{bmatrix}
M_{\Omega_d} & 0 & -\Delta_\Omega + \dfrac{\partial}{\partial n}\bigg|_{\Gamma_c} & 0 \\[2ex]
0 & \lambda F_{\Gamma_c} & -\dfrac{\partial}{\partial n}\bigg|_{\Gamma_c} & I_{\Gamma_c} \\[2ex]
-\Delta_\Omega & 0 & 0 & 0 \\[1ex]
0 & c\Pi_{\mathcal{A}_k} & 0 & \Pi_{I_k}
\end{bmatrix}, \tag{3.86}
$$

and the terms $\delta \boldsymbol{u}_{k+1}$ and $f(\boldsymbol{u}_k)$ read

$$
\delta \boldsymbol{u}_{k+1} = \begin{bmatrix}
\delta u_{k+1} \\
\delta q_{k+1} \\
\delta \lambda_{k+1} \\
\delta \mu_{k+1}
\end{bmatrix} = \begin{bmatrix}
u_{k+1} - u_k \\
q_{k+1} - q_k \\
\lambda_{k+1} - \lambda_k \\
\mu_{k+1} - \mu_k
\end{bmatrix}, \tag{3.87}
$$

$$
f(\boldsymbol{u}_k) = - \begin{bmatrix}
M_{\Omega_d}(u_k - u_d) - \Delta_\Omega \lambda_k + \dfrac{\partial \lambda_k}{\partial n}\bigg|_{\Gamma_c} \\[2ex]
\alpha(M_{\Gamma_c} + \Delta_{\Gamma_c})q_k - \dfrac{\partial \lambda_k}{\partial n}\bigg|_{\Gamma_c} + I_{\Gamma_c}\mu_k \\[2ex]
-\Delta_\Omega u_k \\[1ex]
c\Pi_{\mathcal{A}_k} q_k + \Pi_{I_k}\mu_k - c(\Pi_{\mathcal{A}_k^b} q_b + \Pi_{\mathcal{A}_k^a} q_a)
\end{bmatrix}, \tag{3.88}
$$

where all the terms added to the right-hand side of (3.83) derive from the definition of $\delta \boldsymbol{u}_{k+1} = \boldsymbol{u}_{k+1} - \boldsymbol{u}_k$. With respect to the case without inequalities, we remark the addition of the variable μ and correspondingly of one row and one column in the matrix. We refer to [104] for the definition of the additional blocks, and in particular the terms $\Pi_{\mathcal{A}_k}$ and Π_{I_k}.

3.4.3   Numerical results

In this section, we present some preliminary results on the application of the presented simple control algorithm with fractional regularization term. We remark that the system is solved with a full-monolithic approach, in contrast to the control problem presented in Chapter 2 where an algorithm based on the splitting between the system for the state and the adjoint equations has been introduced. Since the state, the control, and the adjoint variables are all solved with the same numerical system, this approach doesn't need a steepest descent algorithm to converge to the optimal solution. We remark that all
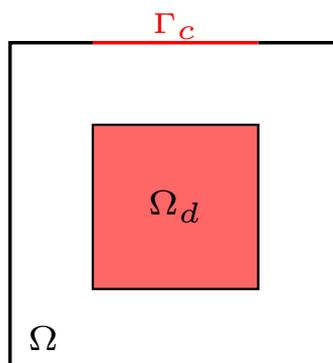
Figure 3.18: Domain Ω with control boundary Γ_c and objective domain Ω_d.

the simulations presented in this Chapter have been carried out using the code FeMTTU introduced above [22].

We report the control problem both with and without the inequality constraints, in order to test all the numerical problems introduced in this section. Unless stated otherwise, we consider the domain in Figure 3.18, where Ω_d is the objective region, Γ_c is the control boundary and homogeneous Dirichlet boundary conditions are imposed on all the boundaries. In particular, $\Omega = \{(x,y) : x \in [0,1], y \in [0,1]\}$ and $\Omega_d = \{(x,y) : x \in [0.25, 0.75], y \in [0.25, 0.75]\}$. It is meshed with a 4×5 grid, that is refined with a multigrid approach. As mentioned above, in this section we solve the following Problem.

Problem 3.3. *Find a state-control pair $(u, q) \in H^1(\Omega) \times Q$ which minimizes the cost functional*

$$\mathcal{J}(u,q) = \frac{1}{2}\|u - u_d\|^2_{L^2(\Omega_d)} + \frac{\alpha}{2}\|q\|^2_{H^s(\Gamma_c)}, \tag{3.89}$$

under the same constraints of Problem 3.2.

We consider different values of the exponent of the Sobolev space of the regularization term s, ranging from 0.25 to 1. Moreover, we introduce the functional

$$\mathcal{L}(u,q) = \frac{1}{2}\|u - u_d\|^2_{L^2(\Omega_d)} + \frac{\alpha}{2}\int_{\Gamma_c} q^2 \, dx + \frac{\beta}{2}\int_{\Gamma_c} \Delta q \, dx, \tag{3.90}$$

where $\alpha = \beta = 0.01$, unless stated otherwise.

The fractional optimal control

We consider now a classical optimal control of the presented problem, without inequality constraints. We test the implemented algorithm with different values of s and for different mesh refinement. For the used grid, the mesh refined 3 times consists of 320 cells and 1353 degree of freedom, the mesh refined 4 times of 1280 cells and 5265 dof, the mesh refined 4 times of 5120 cells and 20769 dof.
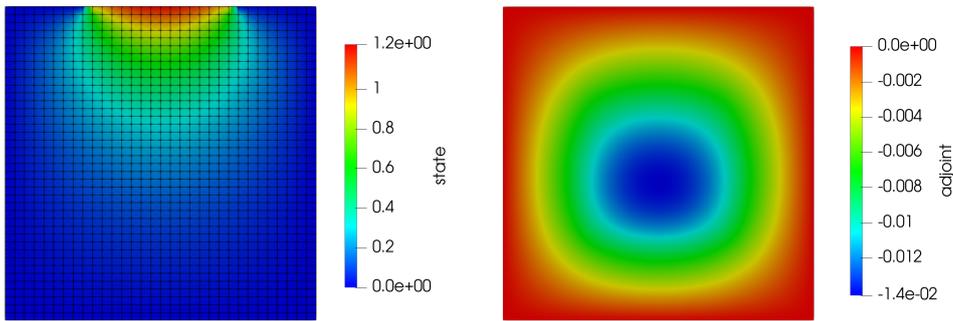
Figure 3.19: State (left, with the used grid) and adjoint (right) fields for the studied case with $s = 0.5$ and 4 grid refinements.

In Figure 3.19 on the left, we report the state field resulting from the numerical resolution of Problem 3.3 with $s = 0.5$ and 4 grid refinements. Note that the state field assumes the value of the control variable on the boundary Γ_c and spreads inside the domain following the constraints of the problem. The control fields on Γ_c depends on the adjoint field, e.g. see in the second line of (3.83). The adjoint field is reported in Figure on the right.

In Table 3.2 we report the numerical results for various s and grid refinements. In particular, we report the values of the global functional $\mathcal{L}$ defined in (3.90) and the distance from the objective calculated as $\frac{1}{2}\|u - u_d\|^2_{L^2(\Omega_d)}$. The global functional, which is minimized by the control algorithm, is greater for low s values. This is due to a lower regularization on the derivatives of the function that affects the global functional. On the contrary, the case with $s = 1$ involves the first derivative in the regularization term, so the Laplacian term in the global functional is minimized. The distance from the objective shows the opposite trend. In fact, the solution is closer to the objective for low values of s. Less regular solutions can control better the behavior of the state variable. Note that the distance from the objective improves with the grid refinement, as expected.

Table 3.2: Global functional and distance from the desired solution as a function of the mesh levels and of the Sobolev exponent s.

s	Global functional $\mathcal{L}(u,q)$			$\frac{1}{2}\|u - u_d\|^2_{L^2(\Omega_d)}$		
	3 Levels	4 Levels	5 Levels	3 Levels	4 Levels	5 Levels
0.25	0.68558	1.47609	3.21692	0.00724	0.00652	0.00650
0.5	0.20364	0.34591	0.60932	0.01392	0.01285	0.01195
0.75	0.08704	0.09048	0.09385	0.01928	0.02067	0.02198
1	0.03017	0.03016	0.03015	0.02914	0.02912	0.02911
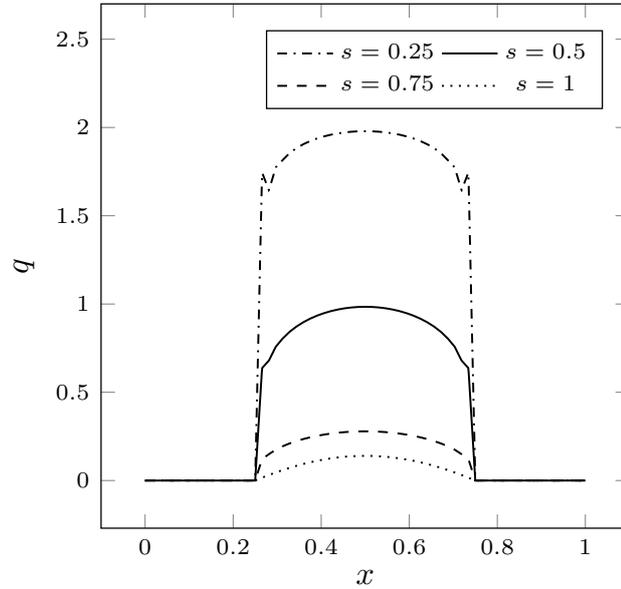
Figure 3.20: Comparison between the control field for 5 levels and various s.

In Figure 3.20 we report the profile of the control variable q on the upper boundary of Figure 3.19. In particular, we plot the variable along the segment between the two points $(0,1)$ and $(1,1)$. The controlled region Γ_c is the segment between the points $(0.25,1)$ and $(0.75,1)$. Note that the control field is strongly influenced by the value of s. In particular, low values of s lead to higher values of the q field, and to less regular solutions, as expected. Therefore, the presented control is strongly influenced by the regularization term.
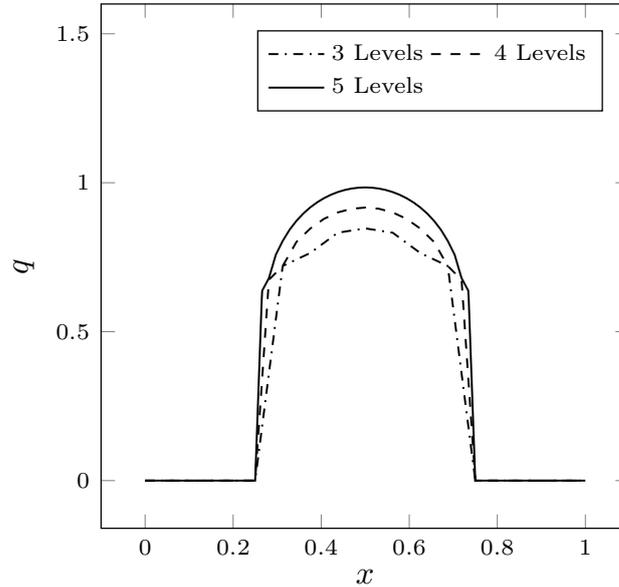
Figure 3.21: Comparison between the control field for $s = 0.5$ and various mesh refinements.

In Figure 3.21 the comparison between different mesh refinements in the case $s = 0.5$ are presented. The solution seems to converge to a more regular one. Moreover, as noted above in Table 3.2, the solution is closer to the desired one for higher mesh refinement.

We also tested the same geometry, with the same parameters but changing the domain $\Omega_d = \{(x, y) : x \in [0, 1], y \in [0.5, 1]\}$ and the desired value $u_d = 1$. The control problem converged in all the tested cases. In particular, in Figure 3.22 the comparison between the q fields for high s values is reported. Again, the control field is strongly influenced by the used considered functional space. Note that for $s \to 1$ the fractional solution converges slowly to the case $s = 1$.

In general, all the tested cases showed good accordance with the desired field and the algorithm always finds a better solution than the initial one. We remark that the results presented in this section are only preliminary results, and further studies and simulations will be developed in future works.

Results with inequality constraints

Lastly, we report some results to show the effectiveness of the formulation of the optimal control based on the inequality constraint, introduced in equations (3.86)-(3.88). In particular, we consider the same numerical setup pre-
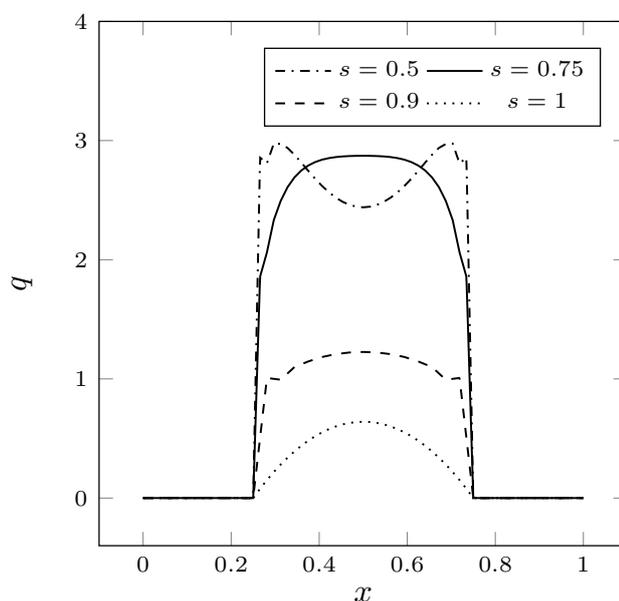
Figure 3.22: Comparison between the control field for 5 levels and various s for $u_d = 1$ and $\Omega_d = \{(x, y) : x \in [0, 1], y \in [0.5, 1]\}$.

sented in the last reported example (i.e. $\Omega_d = \{(x, y) : x \in [0, 1], y \in [0.5, 1]\}$ and desired value $u_d = 1$), and we require an inequality constraint $q \leq 1.5$. We do not impose any lower bound of the solution.
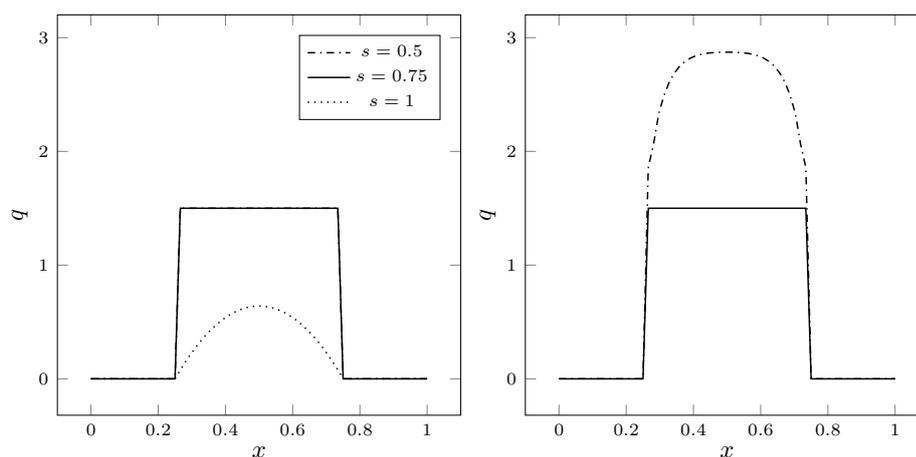
Figure 3.23: On the left, results with inequality constraint $q \leq 1.5$ for various s. On the right, the comparison for $s = 0.75$ between the case with inequalities (continuous line) and the case without inequalities (dashed line).

In Figure 3.23 on the left, we report the result of the presented problem for three different values of s. In particular, the cases for $s = 0.5$, $s = 0.75$ and $s = 1$ are reported. All the reported results was simulated with 5 refinement levels. Note that the solution for $s = 0.5$ and $s = 0.75$ are both limited by the inequality constraint (see Figure 3.22), therefore their profiles are practically identical. On the other hand, the profile for $s = 1$ is not influenced by the inequality constraint, thus it is equal to the case without inequalities. In Figure on the right, we report the comparison between the case with and without the inequality constraint, both with $s = 0.75$. Note that the implemented constraint cut the solution for $q = 1.5$.

We remark that all the results presented on the control with fractional operators are only preliminary, and in future works the application of this technique to more complex physical problems will be studied.

3.5 Fractional operators in quasi-geostrophic problems

In this section, we present a physical application of the fractional Laplacian introduced in this chapter. In particular, we study the quasi-geostrophic models, that has been successfully used for the study of oceanic and atmospheric dynamics in the mid-to-high latitude region of the Earth where the Coriolis effect is significant. The three-dimensional model has been introduced by Charney [105], and many papers and books have been published on this topic over the years. Interested reader can see [106, 107] and references therein. Some works can be found in the literature on the numerical simulations of the quasi-geostrophic flows, e.g. in [108] is presented the numerical simulation of quasi-geostrophic flows using a spectral fractional Laplacian approach.

The main property of the quasi-geostrophic model is the assumption that the potential vorticity is uniform along with the geostrophic flow [109]. Therefore, given a surface $\hat{\mathcal{S}} = (x_1, x_2, x_3) \in \mathbb{R}^3$ such that x_3 is constant, the vorticity $\hat{\psi}(x_1, x_2, x_3, t)$ evaluated above $\hat{\mathcal{S}}$ satisfies

$$\Delta \hat{\psi} = 0 \qquad \text{for } x_3 > 0. \tag{3.91}$$

Moreover, we have $\lim_{x_3 \to \infty} \hat{\psi}(x_1, x_2, x_3, t) = 0$. We can also define the potential temperature on $\hat{\mathcal{S}}$ as $\theta = \partial_{x_3} \hat{\psi}|_{x_3=0}$.

By restricting the domain to a finite, two-dimensional surface and considering the potential vorticity on the surface $\psi(x_1, x_2, t) = \hat{\psi}(x_1, x_2, 0, t)$ and the potential temperature $\theta(x_1, x_2, t)$ conservation, the model reduces to the surface quasi-geostrophic equation [110]. When restricted to the finite surface $\mathcal{S}$, the model corresponds to the elliptic partial differential equation

$$(-\Delta)^{\frac{1}{2}}\psi = \theta \, . \tag{3.92}$$

We use the Riesz fractional Laplacian introduced in section 3.1 to develop a quasi-geostrophic numerical model.

In order to derive the SQG equation, we recall Navier-Stokes equations for an incompressible and inviscid fluid under gravitational and Coriolis body force. The quasi-geostrophic assumption provides that the x_3 component of the velocity vector is negligible with respect to the x_1 and x_2 components. The SQG equations are usually written in terms of two characteristic quantities such as the geopotential $\Psi(\boldsymbol{x}, t)$ and the potential temperature $\theta(\boldsymbol{x}, t)$. The geopotential Ψ is the work done by the gravitational field necessary to lift a unit mass of air upwards at a given height h. The potential temperature θ is the temperature of a unit mass of air at pressure p if it were brought, by adiabatic displacement, to a standard reference pressure p_0.

Under these hypotheses, it is possible to write the velocity vector $\boldsymbol{u}$ in terms of the geopotential such as

$$\boldsymbol{u} = \left(-\frac{\partial \Psi}{\partial y}, \frac{\partial \Psi}{\partial x} \right) = \nabla^\perp \Psi \, . \tag{3.93}$$

This allows us to study SQG systems using only geopotential and potential temperature. If we now consider the boundary restriction over the surface Ω of the quantities, we obtain the SQG system

$$\begin{aligned} &\frac{\partial \theta}{\partial t} + \boldsymbol{u} \cdot \nabla \theta + \kappa(-\Delta)^s \theta = 0 \, , \\ &(-\Delta)^{\frac{1}{2}}\Psi = -\theta \, , \quad \boldsymbol{u} = \nabla^\perp \Psi \end{aligned} \tag{3.94}$$

where $s \in (0, 1)$ and $0 < \kappa \ll 1$ is a dissipation coefficient.

3.5.1 Quasi-geostrophic numerical results

In this section, we present the numerical results obtained from the numerical implementation of the system (3.94) in the presented finite element library

FEMuS. The Riesz fractional Laplacian has been implemented and used to model all the fractional operators in the system. In order to show the robustness of the implemented algorithm, we test the physical model with various initial conditions. In all the presented simulations, we consider homogeneous Dirichlet boundary conditions on all the boundaries. The numerical simulations have been carried out on a domain $\Omega = \{x_1 \in [-\pi, \pi], x_2 \in [-\pi, \pi]\}$, meshed with 1325 quadratic elements (5401 dof). We also consider $\kappa = 0.001$, $s = 0.5$ and a time step of $0.05s$.

We define now some physical quantities useful to evaluate the correct numerical simulation of the quasi-geostrophic flows. In particular, we define the kinetic energy $\mathcal{K}(\theta)$ as

$$\mathcal{K}(\theta) = \frac{1}{2} \int_\Omega \theta^2(\boldsymbol{x}, t) \, d\boldsymbol{x}, \tag{3.95}$$

and the helicity $\mathcal{H}(\theta)$ as

$$\mathcal{H}(\theta) = -\int_\Omega \Psi(\boldsymbol{x}, t)\theta(\boldsymbol{x}, t) \, d\boldsymbol{x}. \tag{3.96}$$
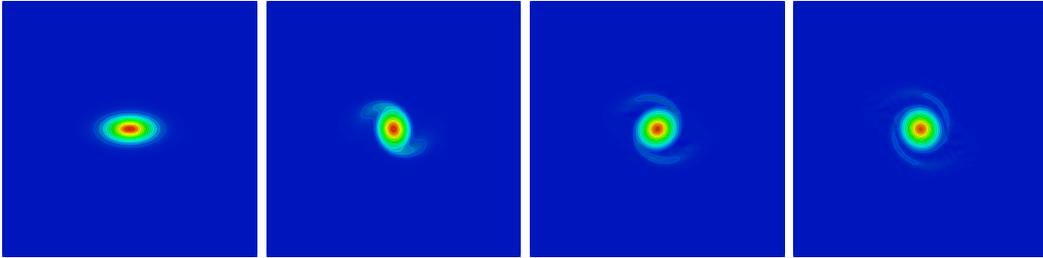
Figure 3.24: Evolution over time of the solution $\theta(x_1, x_2, t)$ with initial condition $\theta_0(x_1, x_2) = e^{-4x_1^2 - 16x_2^2}$ for (from left to right) $t = 0s$, $t = 6s$, $t = 10s$, $t = 15s$.

We first consider the case of a single vortex, resembling cyclonic circulations within the atmosphere. Thus, we consider the initial condition $\theta_0(x_1, x_2) = e^{-4x_1^2 - 16x_2^2}$. In the following, we refer to this case as IC$_1$. In Figure 3.24 the time evolution of the field θ under the introduced initial condition is reported, for $t = 0s$, $t = 5s$, $t = 10s$ and $t = 15s$. Note that, due to the small eccentricity of the imposed initial condition, the vortex does not generate filaments or secondary vortices. The single vortex tends to collapse into a circular configuration.
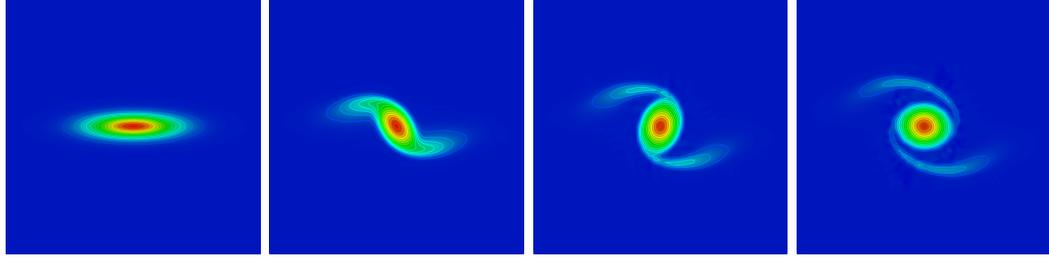
Figure 3.25: Evolution over time of the solution $\theta(x_1, x_2, t)$ with initial condition $\theta_0(x_1, x_2) = e^{-4x_1^2 - 16x_2^2}$ for $t = 0s$, $t = 5s$, $t = 10s$, $t = 15s$.

Then, we consider single vortex with a higher eccentricity, by imposing the initial condition $\theta_0(x_1, x_2) = e^{-4x_1^2 - 16x_2^2}$. In the following, we refer to this case as IC_2. In Figure 3.25, we report the time evolution of the imposed vortex, at the same time steps reported above. Note that the buoyancy turns into a spinning vortex with thin filaments which tend to move away from the vortex center. This behavior is in line with the expectations.
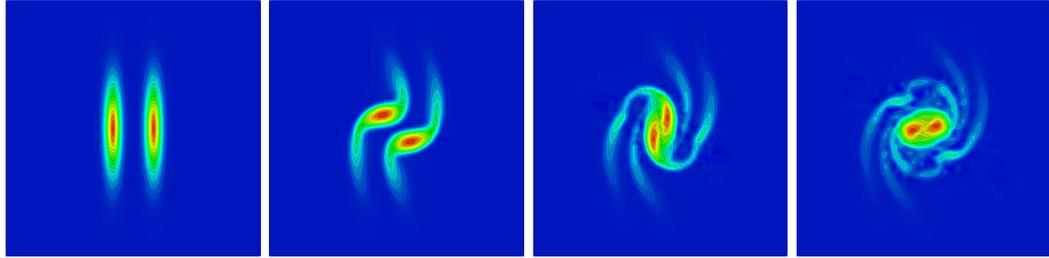
Figure 3.26: Evolution over time for $t = 0s$, $t = 5s$, $t = 10s$, $t = 15s$ of $\theta(x_1, x_2, t)$ with initial condition $\theta_0(x_1, x_2) = e^{-36(x_1 - 0.5)^2 - x_2^2} + e^{-36(x_1 + 0.5)^2 - x_2^2}$.

The interaction between two different vortices, by considering the double Gaussian initial condition $\theta_0(x_1, x_2) = e^{-36(x_1 - 0.5)^2 - x_2^2} + e^{-36(x_1 + 0.5)^2 - x_2^2}$, have also been simulated. In the following, we refer to this case as IC_3. In Figure 3.26, the time evolution of the interaction between the two vortices is reported. Initially, the two imposed gaussian fields do not influence each other. The two fields rotate as introduced in the cases IC_1 and IC_2, and after $t = 5s$ the two Gaussians start to interact. From $t = 10s$ and $t = 15s$, it can be noted that the two vortices tends to merge in a unique vortex. Note that a sharp layer develops between the two vortices. However, although the intensity of the layer is reducing over time, the vortices do not fully merge.

We compare now the three presented initial conditions in terms of helicity
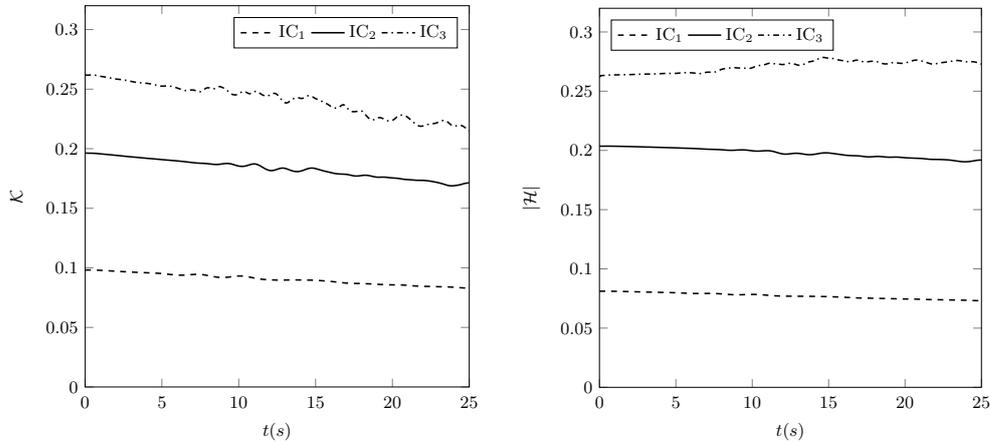
Figure 3.27: Kinetic energy $\mathcal{K}$ (left) and helicity $\mathcal{H}$ (right) in the three presented cases.

and kinetic energy. In Figure 3.27 the kinetic energy and the helicity in the three presented cases are reported as a function of the time. The first two cases, IC_1 and IC_2, present similar, linear behavior. However, both the kinetic energy and the absolute value of the helicity show higher values when the initial condition characterized by the higher eccentricity is used. This result is expected since the vortex generated from the rotation of the ellipse IC_2 generates a greater vortex. The double vortex simulation shows higher values of $\mathcal{K}$ and of $|\mathcal{H}|$, with a less regular evolution over time caused by the interaction between the two vortices.

3.5.2   Numerical optimization of the SQG system

The numerical simulation of the SQG problems entails high computational costs. In fact, the system (3.94) provides the resolution of two fractional Laplacian when $s = 0.5$. Thus, it is important to develop a scalable parallel algorithm and to implement some numerical techniques to reduce the requested computational cost.

The implementation of a parallel code for nonlocal assembly algorithms is not a straightforward task. This is related to the fact that every processor needs information from all the other processors. However, as reported in the previous sections of this chapter, we implemented a parallel scalable algorithm for fractional simulations. In order to further improve the speed of the algorithm, we introduce now a simple technique based on a predictor-

corrector method.

Given a generic time step n, we approximate the solution θ^{n+1} at the step $n+1$ by using an intermediate step $\theta^{n+\frac{1}{2}}$ and a predictor-corrector scheme defined as

$$\theta^{n+\frac{1}{2}} = \theta^n - \delta t (-\Delta)^s \theta^n \,,$$
$$(-\Delta)^{\frac{1}{2}} \Psi^{n+\frac{1}{2}} = \theta^{n+\frac{1}{2}} \,, \qquad \boldsymbol{u}^{n+\frac{1}{2}} = \nabla^{\perp} \Psi^{n+\frac{1}{2}} \,, \qquad (3.97)$$
$$\theta^{n+1} + \delta t \boldsymbol{u}^{n+\frac{1}{2}} \nabla \theta^{n+1} = \theta^{n+\frac{1}{2}} \,.$$

Therefore, by using the system (3.97) instead of (3.94), we can compute only one fractional Laplacian rather than two. This turns into a high reduction of the computational cost of the whole algorithm.
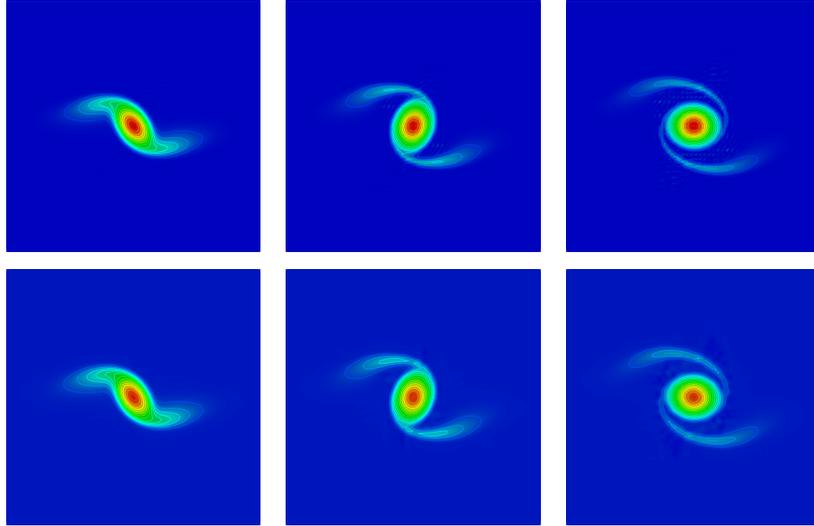
Figure 3.28: Comparison of the evolution over time of the solution with predictor-corrector (top) and the solution with standard technique for (from left to right) $t = 5s$, $t = 10s$, $t = 15s$. The considered initial condition is $\theta_0(x_1, x_2) = e^{-4x_1^2 - 16x_2^2}$.

In order to test the developed predictor-corrector algorithm, we compare the case IC$_2$ introduced above (initial condition $\theta_0(x_1, x_2) = e^{-4x_1^2 - 16x_2^2}$) with the same initial condition used with the predictor-corrector technique. In Figure 3.28, the two cases are compared at three different times, showing qualitatively similar behavior. In particular, the rotation velocity of the vortex is the same in the two studied cases.

The presented results show the robustness of the implemented algorithm and the results are consistent with the expectations. In future works, we will
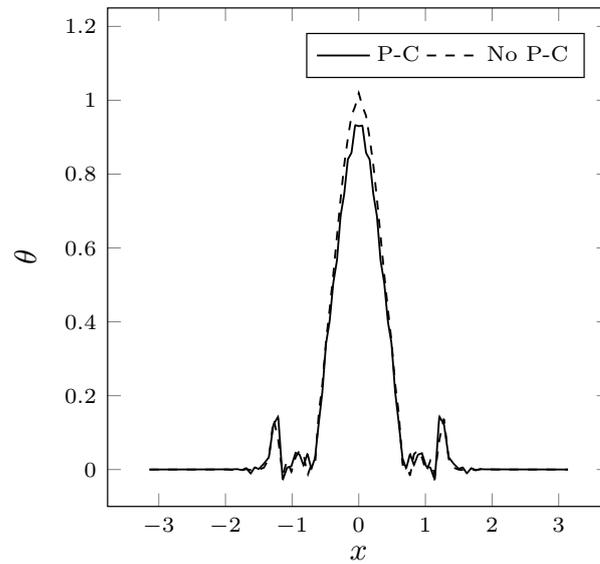
Figure 3.29: Comparison between the case with (P-C) and without predictor-corrector. The θ field along the axis $x_1 = 0$ and $t = 12.5s$ is reported.

further analyze the geostrophic flows and more complex numerical results will be carried out.

Conclusions

In this work, we have applied the optimal control principles to a stationary fluid-structure interaction system based on Koiter shell equations. We first have presented the mathematical model of Koiter's shell equations with and without absorbing boundary conditions. The Koiter model has allowed us to reduce the dimensionality of the solid and the computational cost of the numerical simulation. Since the fluid-structure simulations often depend on pressure waves, we have implemented consistent absorbing boundary conditions to simulate properly the outflow boundary. Numerical benchmarks of the Koiter model do not exist. However, all the presented results have shown good accordance with all the cases in the literature on this topic.

The adjoint-based optimal control theory has been introduced. We have presented a simple mathematical problem and, then, we have applied the control theory to the Koiter fluid-structure interaction problem. In particular, a boundary optimal control of the steady Koiter's model has been introduced from a mathematical point of view, demonstrating the existence of the solution of the mathematical model and the existence of the optimal solution. Moreover, regularity and differentiability properties have been proved. The implemented algorithm minimizes the distance between the solution and a requested objective. To solve the minimization problem we have adopted the Lagrangian multiplier method and the optimality system has been recovered by imposing the first-order necessary conditions. The optimality system has been solved with an iterative gradient-based algorithm implemented in the FEM code. We have implemented an algorithm to find the optimal stress

on a boundary that coincides with the pressure when the tangent velocity is null, and forces the desired displacement on a certain portion of the domain. Numerical results have shown good accordance with the objective in various cases, showing a good convergence with the grid.

In the last chapter, we have introduced the fractional operators. Since the fractional operators and, in particular, the fractional Laplacian have a nonlocal behavior, the numerical implementation is not straightforward. We have presented a classical approach to the fractional Laplacian (called Riesz method) and a spectral approach (spectral and integral methods). We have implemented all the cited methods in a finite element based code, and all the results are in accordance with the analytical results and the numerical benchmarks. Moreover, we have applied the Riesz fractional Laplacian modeling to a simple optimal control problem, where the fractional operator has been used as regularization term. Numerical results for classical and inequality-constrained optimal control problems are shown. We also applied the Riesz fractional Laplacian to quasi-geostrophic simulations. The presented numerical results show good agreement with the expectations, and a numerical scheme based on a predictor-corrector technique has been developed to reduce the computational cost of the simulations.

All the presented numerical algorithms and mathematical schemes can be seen as an improvement of the known optimal control theory applied to the multi-scale FSI model. In future works, the fractional Laplacian approach will be extended from the simple presented mathematical problem to the optimal control of the Koiter fluid-structure interaction.

List of Figures

Bibliography

[1] S. Turek and J. Hron, "Proposal for numerical benchmarking of fluid-structure interaction between an elastic object and laminar incompressible flow," in *Fluid-structure interaction*, pp. 371–385, Springer, 2006. 1, 5, 50

[2] L. Formaggia, A. Quarteroni, and A. Veneziani, *Cardiovascular Mathematics: Modeling and simulation of the circulatory system*, vol. 1. Springer Science & Business Media, 2010. 1, 5, 41

[3] Y. Bazilevs, K. Takizawa, and T. E. Tezduyar, *Computational fluid-structure interaction: methods and applications*. John Wiley & Sons, 2013. 1, 5

[4] W. Koiter, "On the mathematical foundation of shell theory," in *Proc. Int. Congr. of Mathematics, Nice*, vol. 3, pp. 123–130, 1970. 1, 7, 34, 78

[5] F. Nobile and C. Vergara, "An effective fluid-structure interaction formulation for vascular dynamics by generalized robin conditions," *SIAM Journal on Scientific Computing*, vol. 30, no. 2, pp. 731–763, 2008. 1, 7, 34, 38, 39, 61, 80

[6] L. Chirco, *On the optimal control of steady fluid structure interaction systems*. PhD thesis, alma, 2020. 2, 8, 11, 49

[7] F. Abergel and R. Temam, "On some control problems in fluid mechanics," _Theoretical and Computational Fluid Dynamics_, vol. 1, no. 6, pp. 303–325, 1990. 2

[8] M. D. Gunzburger and S. Manservisi, "The velocity tracking problem for navier–stokes flows with bounded distributed controls," _SIAM Journal on Control and Optimization_, vol. 37, no. 6, pp. 1913–1945, 1999. 2

[9] M. D. Gunzburger, L. Hou, and T. P. Svobodny, "Analysis and finite element approximation of optimal control problems for the stationary navier-stokes equations with distributed and neumann controls," _Mathematics of Computation_, vol. 57, no. 195, pp. 123–151, 1991. 2, 84

[10] T. Richter and T. Wick, "Optimal control and parameter estimation for stationary fluid-structure interaction problems," _SIAM Journal on Scientific Computing_, vol. 35, no. 5, pp. B1085–B1104, 2013. 2

[11] L. Chirco and S. Manservisi, "On the optimal control of stationary fluid–structure interaction systems," _Fluids_, vol. 5, no. 3, p. 144, 2020. 2, 70

[12] M. D'Elia, Q. Du, C. Glusa, M. Gunzburger, X. Tian, and Z. Zhou, "Numerical methods for nonlocal and fractional models," _arXiv preprint arXiv:2002.01401_, 2020. 2, 133

[13] E. Di Nezza, G. Palatucci, and E. Valdinoci, "Hitchhiker's guide to the fractional sobolev spaces," _arXiv preprint arXiv:1104.4345_, 2011. 2, 126, 127, 128

[14] A. Lischke, G. Pang, M. Gulian, F. Song, C. Glusa, X. Zheng, Z. Mao, W. Cai, M. M. Meerschaert, M. Ainsworth, _et al._, "What is the fractional laplacian? a comparative review with new results," _Journal of Computational Physics_, vol. 404, p. 109009, 2020. 3, 130, 134, 148, 150, 152, 154

[15] M. D'Elia and M. Gunzburger, "The fractional laplacian operator on bounded domains as a special case of the nonlocal diffusion operator," _Computers & Mathematics with Applications_, vol. 66, no. 7, pp. 1245–1260, 2013. 3, 129, 131, 148

[16] A. Bonito, W. Lei, and J. E. Pasciak, "On sinc quadrature approximations of fractional powers of regularly accretive operators," *Journal of Numerical Mathematics*, vol. 27, no. 2, pp. 57–68, 2019. 3, 137, 141, 142

[17] A. Bonito, J. P. Borthagaray, R. H. Nochetto, E. Otárola, and A. J. Salgado, "Numerical methods for fractional diffusion," *Computing and Visualization in Science*, vol. 19, no. 5-6, pp. 19–46, 2018. 3, 130, 137, 139, 142

[18] "Code FEMuS." https://github.com/FemusPlatform/femus.git. 3, 8, 50

[19] "Numeric platform." https://github.com/FemusPlatform/NumericPlatform.git. 3

[20] G. Bornia, *Analysis of optimal control problems for the incompressible MHD equations and implementation in a finite element multiphysics code.* PhD thesis, alma, 2012. 3

[21] R. Da Già, *Development of a computational platform for the simulation of low Prandtl number turbulent flows.* PhD thesis, alma, 2019. 3

[22] "Code FEMuS (Texas Tech University)." https://github.com/FeMTTU/femus.git. 3, 148, 167

[23] P. Le Tallec and J. Mouro, "Fluid structure interaction with large structural displacements," *Computer methods in applied mechanics and engineering*, vol. 190, no. 24-25, pp. 3039–3067, 2001. 5

[24] E. Aulisa, A. Cervone, S. Manservisi, and P. Seshaiyer, "A multi-level domain decomposition approach for studying coupled flow applications," *Communications in Computational Physics*, vol. 6, no. 2, p. 319, 2009. 5

[25] E. Aulisa, S. Bna, and G. Bornia, "A monolithic ale newton–krylov solver with multigrid-richardson–schwarz preconditioning for incompressible fluid-structure interaction," *Computers & Fluids*, vol. 174, pp. 213–228, 2018. 5

[26] T. E. Tezduyar, S. Sathe, R. Keedy, and K. Stein, "Space–time finite element techniques for computation of fluid–structure interactions," *Computer methods in applied mechanics and engineering*, vol. 195, no. 17-18, pp. 2002–2027, 2006. 7

[27] S. Bnà, S. Manservisi, and E. Aulisa, "A multilevel domain decomposition solver for monolithic fluid-structure interaction problems," in *AIP Conference Proceedings*, vol. 1558, pp. 871–874, American Institute of Physics, 2013. 7

[28] https://github.com/FemusPlatform/NumericPlatform.git. 8

[29] B. S. Kirk, J. W. Peterson, R. H. Stogner, and G. F. Carey, "libmesh: a c++ library for parallel adaptive mesh refinement/coarsening simulations," *Engineering with Computers*, vol. 22, no. 3-4, pp. 237–254, 2006. 8

[30] "PETSc library." https://www.mcs.anl.gov/petsc/. 8, 133

[31] S. Balay, S. Abhyankar, M. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, A. Dener, V. Eijkhout, W. Gropp, *et al.*, "Petsc users manual," 2019. 8, 147

[32] P. G. Ciarlet, *Mathematical Elasticity: Volume I: three-dimensional elasticity.* North-Holland, 1988. 13

[33] T. J. Hughes, W. K. Liu, and T. K. Zimmermann, "Lagrangian-eulerian finite element formulation for incompressible viscous flows," *Computer methods in applied mechanics and engineering*, vol. 29, no. 3, pp. 329–349, 1981. 15, 41

[34] J. Donea, S. Giuliani, and J.-P. Halleux, "An arbitrary lagrangian-eulerian finite element method for transient dynamic fluid-structure interactions," *Computer methods in applied mechanics and engineering*, vol. 33, no. 1-3, pp. 689–723, 1982. 15

[35] P. Lidström, "Moving regions in euclidean space and reynolds' transport theorem," *Mathematics and mechanics of solids*, vol. 16, no. 4, pp. 366–380, 2011. 17, 18

[36] Y. Choquet-Bruhat, C. DeWitt-Morette, C. de Witt, M. D. Bleick, and M. Dillard-Bleick, *Analysis.* Gulf Professional Publishing, 1982. 17

[37] J. N. Reddy and D. K. Gartling, *The finite element method in heat transfer and fluid dynamics*. CRC press, 2010. 19

[38] J. C. Simo and T. J. Hughes, *Computational inelasticity*, vol. 7. Springer Science & Business Media, 2006. 23

[39] S. Bna, *Multilevel domain decomposition algorithms for monolithic fluid-structure interaction problems with application to haemodynamics*. PhD thesis, alma, 2014. 25

[40] R. W. Ogden, *Non-linear elastic deformations*. Courier Corporation, 1997. 25

[41] R. Rivlin, "Large elastic deformations of isotropic materials. i. fundamental concepts," *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 240, no. 822, pp. 459–490, 1948. 26

[42] P. G. Ciarlet, *The finite element method for elliptic problems*. SIAM, 2002. 29

[43] V. Girault and P.-A. Raviart, *Finite element methods for Navier-Stokes equations: theory and algorithms*, vol. 5. Springer Science & Business Media, 2012. 29, 83, 84, 86, 87, 101, 160

[44] R. A. Adams and J. J. Fournier, *Sobolev spaces*. Elsevier, 2003. 29, 84

[45] I. Babuška, "The finite element method with lagrangian multipliers," *Numerische Mathematik*, vol. 20, no. 3, pp. 179–192, 1973. 29

[46] S. Brenner and R. Scott, *The mathematical theory of finite element methods*, vol. 15. Springer Science & Business Media, 2007. 29

[47] D. Cerroni, S. Manservisi, and F. Menghini, "A new moving mesh technique for monolithic multigrid fluid-structure interaction solver," *Recent Advances in Civil Engineering and Mechanics*, vol. 1, pp. 146–152, 2014. 30

[48] J.-L. Guermond, P. Minev, and J. Shen, "An overview of projection methods for incompressible flows," *Computer methods in applied mechanics and engineering*, vol. 195, no. 44-47, pp. 6011–6045, 2006. 34

[49] E. Aulisa, G. Bornia, and S. Manservisi, "Boundary control problems in convective heat transfer with lifting function approach and multigrid vanka-type solvers," *Communications in Computational Physics*, vol. 18, no. 3, pp. 621–649, 2015. 34, 161

[50] D. Cerroni and S. Manservisi, "A penalty-projection algorithm for a monolithic fluid-structure interaction solver," *Journal of Computational Physics*, vol. 313, pp. 13–30, 2016. 34

[51] P. G. Ciarlet and C. Mardare, "An introduction to shell theory," *Differential geometry: theory and applications*, vol. 9, pp. 94–184, 2008. 34, 37

[52] S. Lang, *Introduction to differentiable manifolds*. Springer Science & Business Media, 2002. 35, 39

[53] J. Bonet and R. D. Wood, *Nonlinear continuum mechanics for finite element analysis*. Cambridge university press, 1997. 39

[54] A. Quarteroni and L. Formaggia, "Mathematical modelling and numerical simulation of the cardiovascular system," *Handbook of numerical analysis*, vol. 12, pp. 3–127, 2004. 41

[55] A. Quarteroni, M. Tuveri, and A. Veneziani, "Computational vascular fluid dynamics: problems, models and methods," *Computing and Visualization in Science*, vol. 2, no. 4, pp. 163–197, 2000. 41

[56] A. Chierici, L. Chirco, V. Giovacchini, S. Manservisi, and G. Santesarti, "A multiscale fluid structure interaction model derived from koiter shell equations," in *Journal of Physics: Conference Series*, vol. 1599, p. 012040, IOP Publishing, 2020. 41

[57] D. Cerroni, S. Manservisi, and F. Menghini, "An improved monolithic multigrid fluid-structure interaction solver with a new moving mesh technique," *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 9, pp. 227–234, 2015. 45, 57, 112

[58] J. G. Heywood, R. Rannacher, and S. Turek, "Artificial boundaries and flux and pressure conditions for the incompressible navier–stokes equations," *International Journal for numerical methods in fluids*, vol. 22, no. 5, pp. 325–352, 1996. 46

[59] L. Formaggia, J.-F. Gerbeau, F. Nobile, and A. Quarteroni, "On the coupling of 3d and 1d navier–stokes equations for flow problems in compliant vessels," *Computer methods in applied mechanics and engineering*, vol. 191, no. 6-7, pp. 561–582, 2001. 48

[60] L. Formaggia, F. Nobile, A. Quarteroni, and A. Veneziani, "Multiscale modelling of the circulatory system: a preliminary analysis," *Computing and visualization in science*, vol. 2, no. 2-3, pp. 75–83, 1999. 48

[61] D. Cerroni, R. Da Vià, and S. Manservisi, "A projection method for coupling two-phase vof and fluid structure interaction simulations," *Journal of Computational Physics*, vol. 354, pp. 646–671, 2018. 49

[62] L. Chirco and S. Manservisi, "An adjoint based pressure boundary optimal control approach for fluid-structure interaction problems," *Computers & Fluids*, vol. 182, pp. 118–127, 2019. 49

[63] A. Chierici, L. Chirco, R. Da Vià, M. Manservisi, and S. Magnaniand, "Distributed optimal control applied to fluid structure interaction problems," in *Journal of Physics: Conference Series*, vol. 1224, p. 012003, IOP Publishing, 2019. 49

[64] M. Bukač, S. Čanić, and B. Muha, "A partitioned scheme for fluid–composite structure interaction problems," *Journal of Computational Physics*, vol. 281, pp. 493–517, 2015. 50, 52, 57

[65] J. H. Bramble, *Multigrid methods*. Routledge, 2018. 57, 112

[66] B. R. Noack, M. Morzynski, and G. Tadmor, *Reduced-order modelling for flow control*, vol. 528. Springer Science & Business Media, 2011. 69

[67] M. D. Gunzburger, *Perspectives in flow control and optimization*. SIAM, 2002. 69, 72

[68] J. Nocedal and S. Wright, *Numerical optimization*. Springer Science & Business Media, 2006. 69

[69] J. Kim and T. R. Bewley, "A linear systems approach to flow control," *Annu. Rev. Fluid Mech.*, vol. 39, pp. 383–417, 2007. 69

[70] Y. Yan and D. E. Keyes, "Smooth and robust solutions for dirichlet boundary control of fluid–solid conjugate heat transfer problems," *Journal of Computational Physics*, vol. 281, pp. 759–786, 2015. 69

[71] M. D. Gunzburger and S. Manservisi, "The velocity tracking problem for navier–stokes flows with boundary control," *SIAM Journal on Control and Optimization*, vol. 39, no. 2, pp. 594–634, 2000. 69, 78, 102

[72] M. Perego, A. Veneziani, and C. Vergara, "A variational approach for estimating the compliance of the cardiovascular tissue: An inverse fluid-structure interaction problem," *SIAM Journal on Scientific Computing*, vol. 33, no. 3, pp. 1181–1211, 2011. 70

[73] M. D. Gunzburger, H. Kim, and S. Manservisi, "On a shape control problem for the stationary navier-stokes equations," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 34, no. 6, pp. 1233–1258, 2000. 74, 78, 93

[74] M. Gunzburger, L. Hou, and T. P. Svobodny, "Analysis and finite element approximation of optimal control problems for the stationary navier-stokes equations with dirichlet controls," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 25, no. 6, pp. 711–748, 1991. 74, 102

[75] L. Armijo, "Minimization of functions having lipschitz continuous first partial derivatives," *Pacific Journal of mathematics*, vol. 16, no. 1, pp. 1–3, 1966. 76, 106

[76] C. G. Broyden, J. E. Dennis Jr, and J. J. Moré, "On the local and superlinear convergence of quasi-newton methods," *IMA Journal of Applied Mathematics*, vol. 12, no. 3, pp. 223–245, 1973. 76

[77] B. Muha and S. Canić, "Existence of a weak solution to a nonlinear fluid–structure interaction problem modeling the flow of an incompressible, viscous fluid in a cylinder with deformable walls," *Archive for rational mechanics and analysis*, vol. 207, no. 3, pp. 919–968, 2013. 78, 80

[78] R. Glowinski and O. Pironneau, "Towards the computation of minimum drag profiles in viscous laminar flow," *Applied Mathematical Modelling*, vol. 1, no. 2, pp. 58–66, 1976. 80

[79] A. Mikelić, G. Guidoboni, and S. Čanić, "Fluid-structure interaction in a pre-stressed tube with thick elastic walls i: the stationary stokes problem," *Networks & Heterogeneous Media*, vol. 2, no. 3, p. 397, 2007. 81

[80] S. Čanić, J. Tambača, G. Guidoboni, A. Mikelić, C. J. Hartley, and D. Rosenstrauch, "Modeling viscoelastic behavior of arterial walls and their interaction with pulsatile blood flow," *SIAM Journal on Applied Mathematics*, vol. 67, no. 1, pp. 164–193, 2006. 81

[81] M. D. Gunzburger and S. L. Hou, "Treating inhomogeneous essential boundary conditions in finite element methods and the calculation of boundary stresses," *SIAM journal on numerical analysis*, vol. 29, no. 2, pp. 390–424, 1992. 84

[82] J. Schauder, "Der fixpunktsatz in funktionalraümen," *Studia Mathematica*, vol. 2, no. 1, pp. 171–180, 1930. 86

[83] T. C. Rebollo and R. Lewandowski, *Mathematical and numerical foundations of turbulence models and applications.* Springer, 2014. 87

[84] V. Maz'ya and J. Rossmann, "Mixed boundary value problems for the stationary navier–stokes system in polyhedral domains," *Archive for rational mechanics and analysis*, vol. 194, no. 2, pp. 669–712, 2009. 87

[85] V. M. Tikhomirov, "Fundamental principles of the theory of extremal problems," *New York*, 1986. 94, 96

[86] J. Sokolowski and J.-P. Zolésio, "Introduction to shape optimization," in *Introduction to shape optimization*, pp. 5–12, Springer, 1992. 105

[87] L. Chirco, A. Chierici, R. Da Già, V. Giovacchini, and S. Manservisi, "Optimal control of the wilcox turbulence model with lifting functions for flow injection and boundary control," in *Journal of Physics: Conference Series*, vol. 1224, p. 012006, IOP Publishing, 2019. 125, 161

[88] V. Maz'ya and T. Shaposhnikova, "On the bourgain, brezis, and mironescu theorem concerning limiting embeddings of fractional sobolev spaces," *Journal of Functional Analysis*, vol. 195, no. 2, pp. 230–238, 2002. 127

[89] S. A. Silling, "Reformulation of elasticity theory for discontinuities and long-range forces," *Journal of the Mechanics and Physics of Solids*, vol. 48, no. 1, pp. 175–209, 2000. 128

[90] P. W. Bates and A. Chmaj, "An integrodifferential model for phase transitions: stationary solutions in higher space dimensions," *Journal of statistical physics*, vol. 95, no. 5-6, pp. 1119–1139, 1999. 128

[91] M. Shlesinger, B. West, and J. Klafter, "Lévy dynamics of enhanced diffusion: Application to turbulence," *Physical Review Letters*, vol. 58, no. 11, p. 1100, 1987. 128

[92] Q. Du and X. Yin, "A conforming dg method for linear nonlocal models with integrable kernels," *Journal of Scientific Computing*, vol. 80, no. 3, pp. 1913–1935, 2019. 130

[93] S. Duo, H. Wang, and Y. Zhang, "A comparative study on nonlocal diffusion operators related to the fractional laplacian," *arXiv preprint arXiv:1711.06916*, 2017. 130

[94] G. Acosta and J. P. Borthagaray, "A fractional laplace equation: regularity of solutions and finite element approximations," *SIAM Journal on Numerical Analysis*, vol. 55, no. 2, pp. 472–495, 2017. 130

[95] E. Aulisa, G. Capodaglio, A. Chierici, and M. D'Elia, "Efficient quadrature rules for finite element discretizations of nonlocal equations," *arXiv preprint arXiv:2101.08825*, 2021. 133

[96] A. Bonito, W. Lei, and J. E. Pasciak, "The approximation of parabolic equations involving fractional powers of elliptic operators," *Journal of Computational and Applied Mathematics*, vol. 315, pp. 32–48, 2017. 137

[97] A. Bonito, W. Lei, and J. E. Pasciak, "Numerical approximation of the integral fractional laplacian," *Numerische Mathematik*, vol. 142, no. 2, pp. 235–278, 2019. 137, 143, 145

[98] A. Lunardi, *Interpolation theory*, vol. 16. Springer, 2018. 140

[99] G. Bornia, M. Gunzburger, and S. Manservisi, "A distributed control approach for the boundary optimal control of the steady mhd equations," *Communications In Computational Physics*, vol. 14, no. 3, pp. 722–752, 2013. 161

[100] G. Bornia and S. Ratnavale, "Different approaches for dirichlet and neumann boundary optimal control," in *AIP Conference Proceedings*, vol. 1978, p. 270006, AIP Publishing LLC, 2018. 161

[101] A. Kunoth, "Adaptive wavelet schemes for an elliptic control problem with dirichlet boundary control," *Numerical Algorithms*, vol. 39, no. 1-3, pp. 199–220, 2005. 162

[102] G. Of, T. X. Phan, and O. Steinbach, "An energy space finite element approach for elliptic dirichlet boundary control problems," *Numerische Mathematik*, vol. 129, no. 4, pp. 723–748, 2015. 162

[103] M. Bergounioux, K. Ito, and K. Kunisch, "Primal-dual strategy for constrained optimal control problems," *SIAM Journal on Control and Optimization*, vol. 37, no. 4, pp. 1176–1194, 1999. 165

[104] F. Tröltzsch, *Optimal control of partial differential equations: theory, methods, and applications*, vol. 112. American Mathematical Soc., 2010. 166

[105] J. G. Charney, "On the scale of atmospheric motions," in *The Atmosphere—A Challenge*, pp. 251–265, Springer, 1990. 172

[106] A. E. Gill, *Atmosphere—ocean dynamics*. Elsevier, 2016. 172

[107] G. Lapeyre and P. Klein, "Dynamics of the upper oceanic layers in terms of surface quasigeostrophy theory," *Journal of physical oceanography*, vol. 36, no. 2, pp. 165–176, 2006. 172

[108] A. Bonito and M. Nazarov, "Numerical simulations of surface quasigeostrophic flows on periodic domains," *SIAM Journal on Scientific Computing*, vol. 43, no. 2, pp. B405–B430, 2021. 172

[109] J. G. Charney, "Geostrophic turbulence," *Journal of the Atmospheric Sciences*, vol. 28, no. 6, pp. 1087–1095, 1971. 172

[110] P. Constantin, A. J. Majda, and E. Tabak, "Formation of strong fronts in the 2-d quasigeostrophic thermal active scalar," *Nonlinearity*, vol. 7, no. 6, p. 1495, 1994. 173